NTNU

Norwegian University of
Science and Technology

Implementation of a Relative Positioning System based on Ultrasound

Erlend Fasting
Alexander Rydland Hansen

Master of Science in Electronics
Submission date: January 2016
Supervisor: Odd Kr. Pettersen, IET
Co-supervisor: Tor Arne Reinen, SINTEF

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

Preface

This work marks the end of our studies at the Norwegian Institute of Science and Technology, and then ultimately the completion of our MSc. degree in electronics at the Department of Electronics and Telecommunications (IET). The masters thesis was performed during the fall semester of 2015 and completed in January 2016.

The topic of the thesis was proposed by Odd K. Østern Pettersen at SINTEF ICT, and the goals and layout of the project was composed in collaboration with supervisor Tor Arne Reinen.

The work has involved mapping of system requirements, design and implementation of a prototype system model and testing of the total system. During the process we have had weekly meetings with our supervisor Tor Arne Reinen to discuss the progress and evaluate the solution in the progress.

Trondheim, 2016-01-31

Erlend Fasting

Trondheim, 2016-01-31

Alexander Rydland Hansen

Acknowledgement

The authors would like to thank NTNU and SINTEF ICT, represented by Odd K. Østern Pettersen, for kindly suggesting this project, and for valuable guidance along the way. His visionary approach to research has been a great inspiration during the work on the project.

Special thanks to senior researcher at SINTEF, Tor Arne Reinen for his contribution as supervisor and continuous support to the very end of the project. His wide perspective and detailed knowledge have been of great help to stay on the right track along the way. Also thanks to senior scientist Knut Grythe at SINTEF for his clarifying advices.

The authors would also like to thank Professor Torbjörn Ekman at NTNU for his advices in the field of communication theory, Tim Cato Netland at the acoustic department for supplying us with equipment and Tore Berg and Tore Landsem at the Engineering workshop for always helping out with the construction of equipment.

A special thanks to Professor Peter Svensson for his enthusiasm, inspiring talks and guidance during our time at NTNU.

As a final note, concluding our studies at the Norwegian Institute of Science and Technology, we would like to thank the our families, friends and loved ones for all their patience, love and support.

Erlend Fasting & Alexander Rydland Hansen

Abstract

Speech communication is essential in human interaction. When working or residing in noisy environments, hearing protectors must be used. This leads to reduced communication abilities. If traditional communications systems are used, the natural ability to localize sounds spatially is reduced, resulting in discomfort and unnatural communication.

To aid natural communication in noisy environments, a spatially aided communication system is proposed. The system is able to localize other participants and use their relative position to emulate a spatial position of the speech in the audio domain. Thus, the users will be able to differentiate between participants in a natural way. The system uses ultrasonic signals to determine the position of other users by defining the relative angle and distance between users. HRTF-filtering is applied to an audio signal transmitted by RF.

The work includes the development of a system model based on the human auditory system, definition of parameters related to the application of the system, and design of the physical aspects constrained by a problem definition provided by the research institution SINTEF. Further, a prototype positioning system is implemented on a real-time platform, and the theory of operation is described.

A range of measurements has been performed to investigate the performance of the system. Different transducers positions have investigated, signal-to-noise ratio has been determined, reflection sensitivity is measured, and large-quantity measurements has been conducted to gain statistical knowledge of performance.

Measurements show that the prototype system functions as predicted. The system accuracy is sufficient according to the requirements regarding general performance. However, the implementation is vulnerable to restraints imposed by the physical system model, especially with relation to the placement of transducers. The artificial head used for prototyping introduces sound reflections affecting the measurements to a large extent. The detection algorithms used are not sophisticated enough to prevent acoustic reflections from disturbing the measurements, and Doppler shift due to user motion is not accounted for.

However, investigations of error sources have been conducted, showing that most issues may be mitigated. A range of proposed solutions is discussed to assess these problems. The general evaluation concludes that the system is feasible for further study and implementation.

Sammendrag

Verbal kommunikasjon er essensielt for menneskelig interaksjon. Ved arbeid eller opphold i støyende omgivelser må hørselsvern benyttes. Dette medfører redusert kommunikasjonsevne. Hvis tradisjonelle kommunikasjonssystemer benyttes, vil den naturlige evnen til å lokalisere romlig lyd reduseres. Dette kan resultere i unaturlig kommunikasjon og redusere komforten.

For å forbedre kommunikasjon i støyende omgivelser presenteres her et kommunikasjonssystem som benytter romlig lyd. Systemet kan lokalisere andre brukere og benytte deres relative posisjon til å emulere romlig lyd for brukeren. På denne måten vil brukerne kunne differensiere mellom ulike deltakere i samtalen på en naturlig måte. Systemet benytter ultralyd til å bestemme posisjonen, definert av relativ vinkel og distanse. HRTF-filter benyttes for gjengi lyd overført med RF-signaler romlig til brukeren.

Dette arbeidet inkluderer utviklingen av en system-modell basert på den menneskelige hørselssansen, definisjon av parametere gitt av systemets applikasjon og design av de fysiske aspektene ved systemet, slik det er begrenset av problemstillingen gitt av forskningsinstitusjonen SINTEF. Videre er det implementert en prototype av posisjonssystemet på en sanntidsplattform, og teorien bak systemet er forklart.

Ulike målinger er foretatt for å undersøke ytelsen til systemet. Forskjellige posisjoner for transdusere er undersøkt, signal-støyforhold er bestemt, sensitivitet for refleksjoner er målt og større sett av målinger er gjennomført for å få statistisk forståelse av ytelsen.

Målingene viser at prototypen fungerer som forventet. Systemets nøyaktighet vedrørende generell ytelse er tilstrekkelig for formålet. Det viser seg imidlertid at implementasjonen er sårbar for begrensninger gitt av den fysiske systemmodellen, spesielt med tanke på plassering av transdusere. Kunsthodet som er benyttet i prototypen medfører refleksjoner som påvirker målingene i stor grad. Deteksjonsalgoritmene er ikke sofistikerte nok til å forhindre deteksjon av refleksjoner, og Doppler-skift på grunn av bevegelse er ikke tatt hensyn til.

Undesøkelser av feilkilder har imidlertid blitt utført, og disse viser at de fleste problemene kan forhindres. En rekke forbedringer blir foreslått. Den generelle evalueringen konkluderer med at systemet er gjennomførbart, og at videre arbeid og implementasjon bør utføres.

Contents

Acronyms

A/D conversion Analogue to Digital conversion

BPSK Binary Phase-Shift Keying, a modulation technique

CC Cross-Correlation

CDMA Code Division Multiple Access, a multiple access communication strategy

DSP Digital Signal Processor

DS-CDMA Direct Sequence Code Division Multiple Access

DSSS Direct Sequence Spread Spectrum

EBU European Broadcast Union

FDMA Frequency Division Multiple Access, a multiple access communication strategy

FFT Fast Fourier Transform, an implementation of the Discrete Fourier Transform

FH-CDMA Frequency-Hopping Code Division Multiple Access

FHSS Frequency-Hopping Spread Spectrum

FPGA Field Programmable Gate Array

GUI Graphical User Interface

IDE Integrated Development Environment

IR Impulse Response

IQ In-phase/Quadrature

MAI Multiple access Interference

PCB Printed Circuit Board

PN-sequence Pseudo-random Noise sequence

PSD Power Spectral Density

RTAP Real Time Audio Processor, a class in the MATLAB implementation

SFG Signal Flow Graph

SPL Sound Pressure Level, ratio referenced to 20μPa

TDMA Time Division Multiple Access, a multiple access communication strategy

TDOA Time Difference Of Arrival

TOF Time Of Flight

USB Universal Serial Bus

Symbol list

$c[n]$ Carrier signal

f_s Sample rate

f_c Carrier rate

f_m Chip rate

L_{signal} The number of samples per modulated signal at f_s

L_{chip} Length of chip sequence

$L_{chip,pre}$ Length of preamble chip sequence

$L_{chip,tot}$ Sum of preamble and unique sequence chip length

$m[n]$ PN chip sequence at chip rate

$\dot{m}[n]$ PN chip sequence at sample rate

$m_r[n]$ Received BPSK modulated

M Chip sequence number

$r[n]$ Received signal

$r_{dm}[n]$ Received down-mixed signal

$r_{LP}[n]$ Received down-mixed and low-pass filtered signal

$s[n]$ Modulated and up-mixed output signal

Δ_k Relative offset between incoming signal and buffer signal inside a buffer, also denoted lag, for chip sequence k

Δ_{dec} Relative offset correction due to decimation decision point

$\hat{\phi}$ Estimated carrier phase for received signal

Υ Cross-correlation coefficient

$\Upsilon_{threshold}$ Detection threshold for cross-correlation coefficient

Chapter 1

Introduction

Speech communication is essential to the human kind. In all sorts of situations, people are required to communicate to solve problems. The information is shared both visually and verbally, but if the speech ability is reduced, the information capacity is drastically hindered. In noisy environments, hearing protection is needed to avoid damage to the auditory system. If speech capability is essential, a communication system is often embedded in the hearing protection. However, traditional systems reduce the functionality of the auditory sense to a large extent. Most importantly, the inherent ability to localize sounds spaciously is eliminated, affecting the capability to separate different voices in an environment. As well, the environment properties are not perceivable, and in total, discomfort and ambiguity may be the result.

Many environments in the industrial world suffer from high noise levels. Typical environments may be of industrial nature, but non-industrial environments are also of concern. Some examples of noise-contaminated environments which require speech communication are; installation and maintenance in industrial plants, security personnel working at concerts and mechanics working in car workshops. The latter environment is often exposed to transient noise, and this may cause the personnel to reject wearing hearing protection as the absence of a fully functioning auditory sense reduces the general communication abilities.

The motivation for this work is to contribute to the development of a spatially aided communication system for noisy environments. This report aims to investigate and describe a positioning system, thus creating a framework for a complete spatially aided system in the future. Positioning should be independent of external framework, such that only a wearable system for each user is required for the system to function. Indeed, this may provide incentives for adoption in the commercial world, providing an easy-to-use solution to the communication problem.

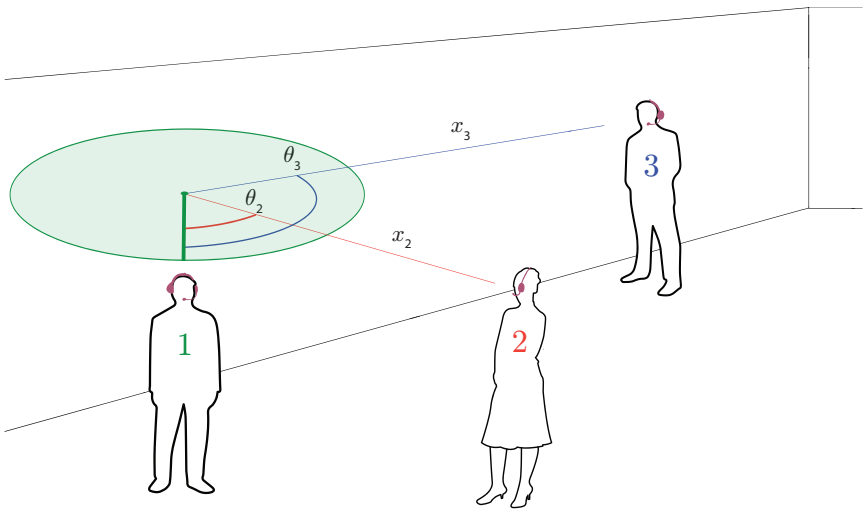


Figure 1.1: Illustration of the relative positioning system for three participating subjects.

1.1 Problem confinement

The benefit of a spatially aided system opposed to a traditional system is that it allows the users to localize the communicating participants in the audio domain. The human auditory system uses a range of cues to locate sound sources, and with the knowledge of the functionality of the auditory system, technology may exploit the cues to generate spatial sound. Technology for emulation of a single channel signal into the spatial domain already exists for headphones. The technology is named 3D-sound and is based on head-related transfer functions (HRTF) which are a general or personal description of the influence the head and ear have on the sound input at each ear for different incident angles.

The contribution of this work will as mentioned be focused on the positioning system. To aid a 3D sound system generate a spatial representation of sound, the most important properties are the relative azimuth angle between the source and the receiver and the distance between them. The altitude difference is also needed to locate the source in a three-dimensional space. In most conditions, the source and receiver may be assumed to be at the same level. Thus, calculating the relative azimuth angle and distance is the primary objective of the proposed positioning system. The azimuth angle may be input directly in HRTF filters, whereas the distance may be used to create a depth perspective. Further exploitation of the distance may enable environmental simulation, for instance to add reverberation to enhance the perception of a real room.

In principle, several positioning technologies may be capable of determining the relative angle and distance between users. Both RF, visual and ultrasonic systems exist. However, considering the application, an ultrasonic system has some

significant advantages:

a) Ultrasound is limited by the same constraints as speech, i.e. physical boundaries of confined space
b) Ultrasound is better suited for short-distance calculations than RF as the speed of sound is much lower than the speed of electromagnetic waves
c) Visual systems are usually more expensive than the two systems mentioned above, and requires considerably higher processing requirements. Thus, these systems are better suited for permanent installation and mainframe processing.

This work is a collaboration between the authors and the research institution SINTEF, and considering possible applications, an ultrasonic positioning system is desired. The task description is formulated in cooperation with SINTEF:

> A system to simplify communication in noisy environments using hearing protection shall be studied. Using ultrasonic sound, distance and angles to other units shall be measured, such that these parameters, in agreement with an inertial measurement unit, may be used to process sound through HRTF-filters. The system shall operate without any other components than the user-worn equipment. Different methods shall be evaluated considering measurement method, noise immunity, the number of communicating units other physical limitations. It is desired that the work results in a working prototype, such that the system may be evaluated subjectively.

In this report, the results of the completed work is presented. All aspects described in the task description has been investigated, and a prototype has been made and tested. Unfortunately, time did not allow for subjective evaluation including HRTF-filters, but the prototype is ready for such tests.

The scope of this work spans several speciality fields. A profound and exhaustive discussion of all elements will thus not be possible, but references to relevant research will be given throughout the text.

An illustration of the principal function of the system is shown in Figure 1.1. The figure shows the measured parameters θ and x, denoting angle and distance respectively, for subjects 2 and 3 related to subject 1. All subjects will measure the same parameters to position any other subjects using the same system.

1.2 Structure of the report

The following report is divided into several chapters, each presenting different aspects of the system. *Chapter 2 - System Design* presents a model of the system, requirements based on the model, and theory of operation for the positioning system.

Relevant parameters for optimal function of the system will be stated. *Chapter 3 - System Implementation* describes the implementation of the prototype. Hardware and the restraints induced by the hardware will be presented, and all important design parameters will be discussed. Details related to the theory of operation for specific system modules will also be presented. Chapter 4 discusses and presents measurements performed with the prototype, and verification of system requirements from Chapter 2 will be addressed. Finally, a conclusion is given in Chapter 5. The rest of this chapter will be dedicated to previous work and existing solutions.

1.3 Previous work

Ultrasonic positioning is well examined in the literature. Several prototype systems have been made, but the focus has been on absolute positioning systems with mainframes and room-mounted transceivers. Some systems are commercially available. Some types of systems, such as the ActiveBat[1] produced by AT&T, utilises a wearable badge that transmits an ID. The signal is received by numerous receivers in the ceiling, and a centralised computer calculates the absolute position by trilateration based on the distances from a transmitter to the receivers. The accuracy is considered good, around 3cm, but requires extensive amounts of installed equipment. The Cricket system[2] calculates the position of the wearable receiver instead, offering a user privacy similar to the GPS system. However, the system still requires infrastructure, as multiple transmitters need to be placed at known positions in the ceiling of every room of operation. Position accuracy of the Cricket system is estimated to around 10cm.

Sonitor[3] is another indoor tracking solution, offering room level accuracy using small tags that can be attached to equipment and persons. A detector in each room sends the tag position to the centralized unit over a standard computer network. Randell and Muller[4] also propose a similar system that allows a wearable tag to automatically calculate its position by utilising transmitters placed in the ceiling of a room. Common for all the mentioned systems is that they all rely on RF transmission between the ultrasound transmitters and receivers to enable synchronization.

The Dolphin system[5] was the first to use a call-response scheme, eliminating the need for RF synchronization. A spread-spectrum signalling technique is also used, allowing multi-user access and identification. Location is calculated by measuring the time-of-flight. A special near-omnidirectional transducer is employed in the Dolphin system. However, the implementation is bulky and requires substantial power, thus not suited for a mobile system. Several improvements has been gained over the Dolphin system. For instance, a system proposed by Gonzales and Bleakley[6] uses frequency hopping spread-spectrum and a MEMS sensor array as the receiver, resulting in a location accuracy of 1cm.

Chapter 2

System Design

This chapter will give a description of the system, both regarding the overall system and the positioning system further investigated in this work. The purpose and function of the system will be discussed, and a system model will be described. Relevant theory regarding the auditory sense will also be presented. Based on the model, requirements for the system will be stated. The theory of operation for the technical aspects of the system, specifically the positioning system, will be presented. Finally, in Section 2.5, the requirements and specifications are summarized.

2.1 System model

The paramount objective of the proposed system is to provide a speech communication system that enables multiple users to conduct natural speech communication in a noisy environment. This section discusses the fundamentals of the auditory system's spatial location abilities and the physical aspects that will affect the model.

2.1.1 Conversations in confined space

When humans communicate verbally, they may easily determine the position of the involved persons. Besides visual impressions, localization is achieved by utilizing only two inputs, the sound pressures at our left- and right- ear. The human auditory system exploits three properties of the sound to localize a source; time/phase offset, level difference and spectral content. By evaluation of the two inputs, direction and distance of the source can be estimated, along with environmental characteristics such as room size, absorption and large nearby objects such as walls.

The mechanisms of the auditory system is thoroughly covered in literature, for example by Blauert[7]. The hearing system is capable of accurately determining displacement of the source for as low as 1° in the azimuth plane under ideal circumstances. This corresponds to a time difference of $\Delta t = 10\mu$s for an average

head. For sound incident at the normal of the median plane, a 10° displacement is required to produce a perceivable change of angle as described in [8]. Figure 2.1 clarifies the definitions for sound incidence.

The most imprecise location performance is observed for sources located in the median plane, as these are solely based on spectral cues. This also applies to sources right behind the subject, which may be difficult to replicate for a general spatial emulation system. However, the importance of visual cues must not be overlooked. The impressions of the two senses are always correlated, and if visual contact is established, audio cues may be downgraded by the visual cues.

A few assumptions can be made about communicating in confined spaces:

1. A person entering a room will normally cause visual attention, caused by sound cues or visual movement related to the entry.
2. The visual attention causes faces to align face to face.
3. The visual attention is often followed by a verbal greeting which determines if a conversation will be conducted.
4. If a conversation is to be conducted between the two, they will probably move to a relative distance that enables communication with proper SNR.

When communicating in noisy environments, the subjects will not be attentive to each other before visual contact is established. If communication is to be conducted, the distance between them that is needed to be able to communicate will be very short, with or without hearing protectors. This will hamper the ability to perform tasks while communicating. If the users have radios available, they will need to communicate via an available channel that is not occupied by others to ensure privacy.

2.1.2 Kinetics

The velocity of a communicating subject is difficult to estimate on a general basis, but for most situations fast walking speed can be assumed as maximum. Thus, $v_{subject,max} = 5\text{km/h} \approx 1.4\text{m/s}$ and the maximum difference will be the double, $v_{diff} \approx 2.8\text{m/s}$. However, in general, the speed difference between two subjects is assumed to be approximately 1m/s.

Regarding the angular velocity, consider a typical situation where a person moves past another with a speed $v = 1\text{m/s}$ at a distance of 2m. The angular velocity will be approximately 28°/s in this case.

The highest angular speeds will be a result of head rotation of the subject itself. Informal tests reveal that head rotation speeds as high as 600°/s may be observed.

Maximum distance for speech communication varies with the environment and situation. For longer conversations, under normal circumstances, people tend to move to a closer distance, between 0.5m to 2m. This is both due to the comfort of visual contact between the subjects and room-specific parameters such as reverberation

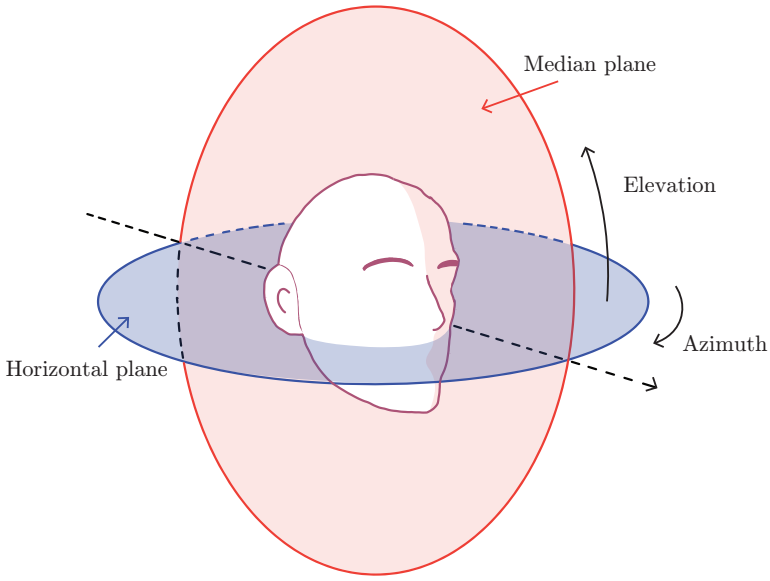


Figure 2.1: Illustration of head-related orientation planes. The frontal plane is excluded for readability.

(the double of the hall radius is considered as a good indicator, increasing SNR for the conversation) and noise. For short conversations, instructions or commands, distances may be much larger, and possibly up to 50m. This normally implies that raised speech levels are used, as the attenuation and reverberation will affect the perception of sound to a large extent.

2.1.3 Multi-participant conversation

For conversations with several participants, the auditory system uses several cues to separate the subjects. An essential role is played by the spatial separation - if the spatial separation fails, only spectral cues and personal speech characteristics will be available for the separation, thus, identification will be difficult. This issue becomes more significant if the subjects involved has not been able to familiarize a priori.

For a traditional speech communication system, the spatial dimension is lost, and speech from all users appears "in-head" as a standard intercom headset delivers all speech as one signal at both ears.

Most conversations with multiple participants are assumed to be limited to around eight subjects. This is the limit for effective discussions in a group. Above this number, mostly briefings from one leading speaker will be the normal case.

Another issue with traditional speech communication systems is privacy. Limitations such as walls or distance do not apply to such systems to the same extent.

2.1.4 The human head

The size and shape of the human head vary between individuals, and it naturally correlates with the person's weight and height. The most significant head characteristic for the model is the width of the head, as this is the parameter that determines the spacing between receivers in the proposed system. An average of 14.5cm for spacing between ears is a decent estimate [9].

2.1.5 Hearing protector

A hearing protector headset is one of the fundamental components of the spatially aided communication system, providing the passive attenuation required for use in noisy environments. The headset may also serve as housing for the hardware implementation of the system. The physical size may vary by type, but a typical addition of 6cm to the head width at each side is a fair estimate.

2.1.6 HRTF spatial emulation

As mentioned in Section 1.1, HRTF processing may be used to emulate the spatial dimensions for a speech source signal. HRTFs can be made by impulse response measurements inside the outer ears of a human subject or an artificial head for all possible angles of incidence around the head.

As HRTFs are implemented as different filters for different angles, or an adjustable filter, the processing introduces a processing delay. According to Brungart et al. [10], latency values of lower than 60 ms are likely to be adequate for virtual audio applications. This number may be a guideline for the update rate of the positioning system.

2.1.7 Noise cancelling

Active noise cancelling is a technique developed to reduce the noise inputs at the ears using opposite phase sound to cancel the incoming noise. The implementation is often based on adaptive filters, with microphones mounted at the ear can as inputs. Active noise cancelling systems may be able to achieve greater reduction in noise levels than passive protection, but it is worth noticing that noise cancelling primarily reduces low-frequency random-process noise. For the recording of a speech signal, a noise cancelling algorithm may also be applied.

2.2 System requirements

The problem description stated in Section 1.1 specifies that the system should be stand-alone and not constrained to a centralized mainframe implementation. An easy-to-use overall system is of preference, and the system should be based on relative positioning. This means that the participants merely needs to quantify the positions of others, related to ones self. The system described in this work is preferably designed to integrate well into a hearing protector. Besides, noise cancelling technology may then be integrated to form total system to comprise both spatial emulation of speech communication, hearing protection and noise cancelling.

To fulfil the above requirements, the system will need to consist of three main parts:

a) An ultrasound positioning system calculating the distance and determining the relative angle to the other subjects.

b) Radio communication unit for transmission of the potentially noise-cancelled speech recordings.

c) HRTF processing of the received voice signals aided by the relative positioning system.

This work will only focus on part a) of the system above, but the design will have to take into consideration aspects of the total system, both regarding physical constraints and performance parameters.

The complete system will consist of several communication nodes. Each user will utilise a node system that will supply the user with the speech from other subjects over radio, processed according to the relative position. Consider a polar plot, depicted in Figure 2.2, where the node of interest is positioned in the centre. The relative position of other nodes will be characterised by two parameters; the distance x, and the relative azimuth angle θ $[-180\,^\circ, 180\,^\circ]$ in the horizontal plane. $0\,^\circ$ is defined as the front of the node.

The elevation in the medial plane will not be quantified in the proposed system. It is assumed that the nodes will be on the same vertical level, and if not, the azimuth position should be sufficient in most conditions.

The relative distance between the nodes may be used for several purposes. For a realistic emulation of the system, the distance may be used to attenuate the signal, add delay and simulate room reverberation. Thus, the system will perform as if the environment is noise-free. Detection of the nodes will not be possible if a node is outside in another room, and this information may be used to increase privacy by attenuating the audio channel completely. Another scenario is to enhance the communication ability in the room. The system may utilise spectral cues, for example, add reverberation or adapted reduction of high-frequency content to simulate air absorption while maintaining a sufficient signal level. This way, an indication of distance will be given, without compromising the necessity of actual communication, which of course is essential in any communication system.

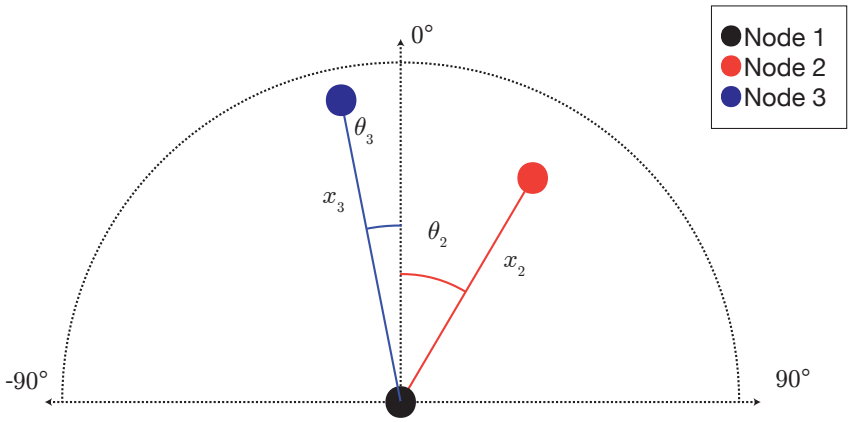


Figure 2.2: System of 3 nodes from the perspective of node 1, plotted for angles $-90°$ to $90°$.

As may be clear to the reader, each node will have to be identified, such that radio communication may be related to the position.

2.2.1 Physical design and user interface

Key elements of the proposed system are ease-of-use and portability. The ultrasonic positioning system and its identification scheme should be integrated with the communication radio such that a radio link is established upon ultrasonic contact. Identities should be automatically assigned. These features should result in a plug-and-play system where no configuration is necessary. As mentioned in Section 2.1.1, specific behaviour due to noisy environments may be exploited for the sake of handshaking and determining the identities of the users in the system.

The portability aspect is an important feature. The system should preferably consist of a single unit to ease the use. Considering the traditional head-worn hearing protector, the positioning system should be physically integrated. The noise cancellation technology mentioned in Section 2.1.7 may as well be integrated to further increase the performance of the hearing protector. Thus, the overall system will benefit from using the same components for multiple subsystems. This may apply to signal processing hardware, but most obvious are sharing of transducers. Therefore, the transducer configuration for the positioning system should take noise cancelling and speech recording into consideration.

2.3 Ultrasonic positioning

As mentioned in Section 1.3, a number of location systems based on ultrasonic signals has been developed and some are currently commercialized. Their performance is well-documented and many ultrasound-specific issues have been investigated and mitigated. The benefits of using ultrasonic signals are the slow propagation speed through the air as opposed to RF signals. The ultrasound transducers are relatively cost efficient and require lower processing speeds due to the lower frequency. However, issues such as multi-path reflections, the directivity of transducers and air absorption will cause challenges.

Ultrasound is characterised as sound above the perceptive limit of the human auditory system. Thereby, it shares the same physical properties as audible sound. The characteristics of audio signals are thoroughly covered in the literature, for example by Krokstad[11], and will not be discussed in detail. However, some key challenges of ultrasonic transmission should be mentioned. The behaviour of waves concerning obstructing surfaces is important. Most important are reflections caused by physical boundaries, and diffraction, the wave-bending effect which appears when sound hits an obstructing object and propagates around it. The behaviour of waves can be solved analytically for simple geometries, such as a solid sphere in a plane wave, which can be comparable to the human head.

Considering an analogy from Krokstad[11] concerning a plane wave incident on a sphere: For the case where the diameter of the sphere(d) is much larger than the wavelength of the incoming plane wave, $d >> \lambda$, phase rotation causes addition of waves on the shadow side of the sphere. An attenuation of the incoming sound wave will thus be the result, and an increasing number of lobes will be apparent around the head for increasing frequency. For the design of an ultrasonic positioning system, this will affect the systems ability to detect ultrasonic signals on the shadow side of the head. The wavelength is given as

$$\lambda = \frac{c_s}{f}. \tag{2.1}$$

The human voice is normally located around 1kHz, resulting in a wavelength $\lambda =$ 340mm. Then consider a pulse centre frequency of 40kHz. The wavelength is thus $\lambda = 8.5$mm. The difference should be obvious; the speech may, as described above, more easily pass around a head than the ultrasonic pulse.

Air absorption, characterised by the absorption coefficient α describes the intensity loss per metre induced by the viscous, heat conduction and internal molecular relaxation processes in the transmitting medium [12]. As can be further investigated in Kinsler and Frey [12, equation 8.5.3], it is directly proportional to the frequency squared of the transmitting sound wave and will therefore reduce the intensity of ultrasonic signals to a larger degree than audible sound waves. Figure 2.3 shows the air absorption coefficient as a function of frequency and relative humidity (%), as described by [13].

The speed of sound varies with environmental conditions like humidity, pressure

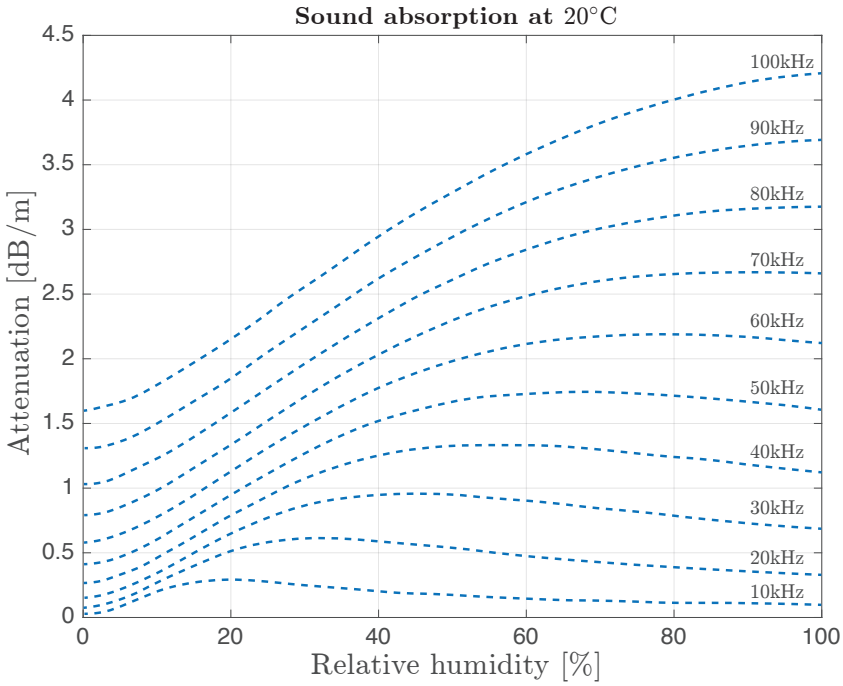


Figure 2.3: Attenuation of ultrasound for different frequencies due to air absorption plotted for relative humidity at 20°C.

and temperature. However, for temperatures around 0°C, the expression for the speed of sound is simplified to

$$c_s = 331.6 + 0.6T \tag{2.2}$$

where T is the temperature in °C.

Due to the low propagation speed of sound as opposed to radio signals, it is expected that relatively small velocities will induce Doppler shift. The Doppler shift caused by a moving source is according to [14] given as

$$f_D = f \pm \frac{v}{c_s} \tag{2.3}$$

where c_s is the speed of sound, v is the relative motion of the source and f is the frequency. Considering a transmission frequency $f = 40$kHz, the Doppler shift induced will be as much as $f_D = \pm 230$Hz depending on the direction of movement.

2.3.1 Estimation of angle

To measure the angle, a pulse needs to be received in at least two positions on the node. The time difference of the pulse reception (TDOA, *Time Difference of Arrival*) may be used for angle calculation. However, the use of only two transducers will introduce an ambiguity; the incoming signal may be interpreted as for instance $-15°$ or $-165°$ as illustrated in Figure 2.4.

Considering the auditory system's ability to estimate incident angle by exploiting only two audio signals, this is a result of a spectral content analysis based on the shape of the outer ears. This would be rather difficult to implement as it requires complex signal analysis and processing, as well as a signal with a very wide spectral content. However, it is worth to notice that ambiguity may be evident also for the auditory sense. The front-back ambiguity is based on the human mind's training of the hearing, and thus very difficult to handle for HRTF systems based on artificial heads.

To conquer the ambiguity, different methods may be applied. For instance, configurations with asymmetric microphone directivity and level calculations could be investigated. This may yield results, but one would have to consider reflections both from the head and surrounding environment and obstacles. Another approach would be to use a third transducer. For instance, if a dedicated microphone placed nearby the user's mouth is to be used for speech recording, the microphone could exploit the shadowing effect resulting from a rear-incident sound source.

The previous example presupposes that the two transducers are spaced on an axis parallel to the ears, thus in the same configuration as the ears. The location is depicted in Figure 2.6. Considering that the auditory system exploits time difference in the same way as proposed here, similar characteristics will be obtained. Assuming a plane wave pulse incident on two spaced transducers, the angle may be calculated as

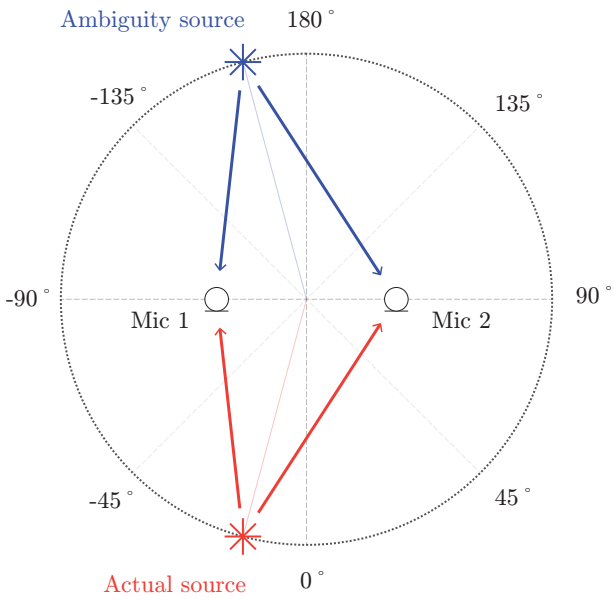


Figure 2.4: Illustration of source-mirroring ambiguity for dual transducer configuration.

$$\theta = \sin^{-1}\left(\frac{\Delta t \cdot c_s}{d}\right) \tag{2.4}$$

where Δt is the time difference of arrival, c_s is the speed of sound and d is the spacing between the microphones. A plot of equation 2.4 is illustrated in Figure 2.6a. The time difference per angle is plotted in 2.6b. Considering the figure, it is clear that for incidence angles approaching 90°, the time difference per angle becomes very small compared to a normal incidence. Thus, the accuracy will become lower as the angle approaches 90° incidence. This corresponds well with the human auditory system's accuracy, as discussed in Section 2.1.1.

As stated above, equation 2.4 assumes a plane wave. Plane waves may be assumed if the source is in the far field. For near-field sources, the equation becomes more complex as the waves are spherical. Far-field is assumed when the distance x between the nodes are large compared to the spacing between transducers d. Thus, for spacings corresponding to the ear spacing, $d = 0.20$m, far-field may be assumed for $x > 2.0$m. The more complex near-field model is compared to the simplified far-field model in Figure B.1 included in the appendix. For the proposed system, the far-field model is considered sufficient, as very close distances $x < 1$m is not expected. Visual cues are expected to be dominant in these situations, and an imprecise angle will not degrade the user's spatial perception.

Selection of the spacing must be seen in relation to the head size and ear cans

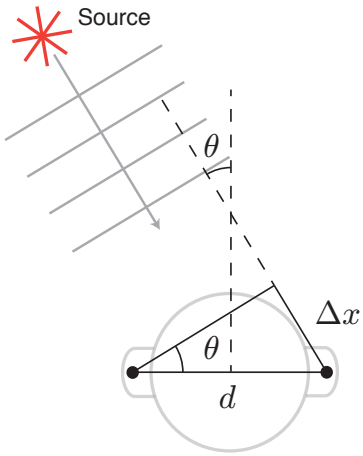
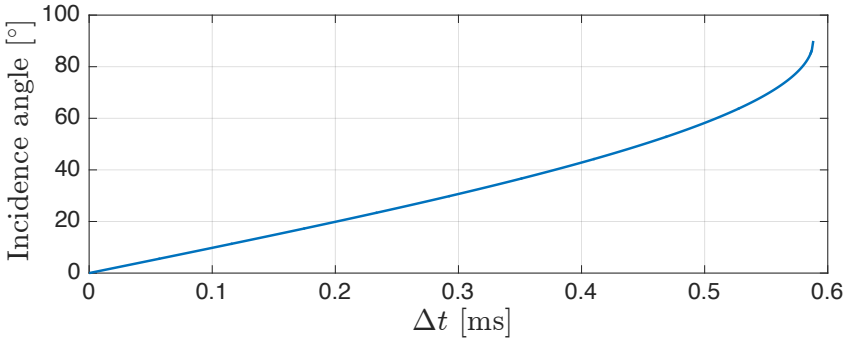
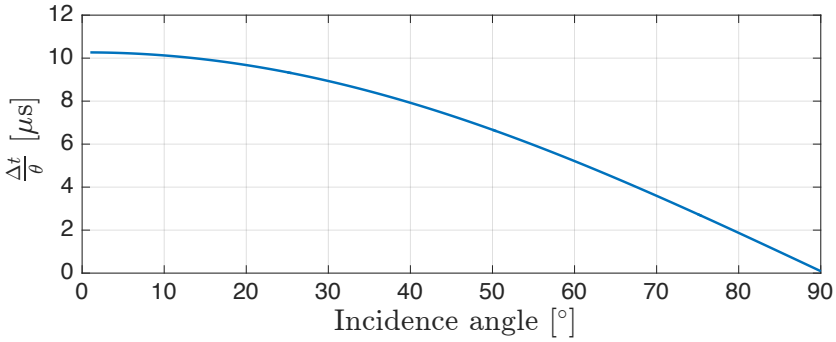


Figure 2.5: Transducer configuration equivalent to the auditory system. Plane wave incidence at angle θ causes an extra travel distance Δx.



(a) Angle of incidence θ as a function of time difference Δt.



(b) Time difference pr degree $\frac{\Delta t}{\theta}$ as a function of incidence angle θ

Figure 2.6: Illustration of incidence angle calculation with spacing $d = 0.2$m and speed of sound $c_s = 340$m/s.

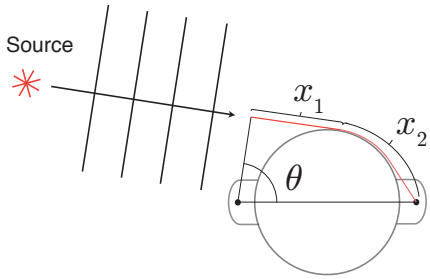


Figure 2.7: Extended travel time at angles larger than line-of-sight.

imposed by the hearing protector. From Section 2.1.4, the average male head's ear spacing is approximately 14.5cm. The ear cans described in Section 2.1.5 spans about 6cm on each side of the head. Thus, a transducer localization on top of the ear cans will provide a spacing d a bit wider than the auditory system. For the prototype, a spacing $d = 0.22$m is taken as the design constraint. Thus, angular interval with the line-of-sight to sources will be a bit larger than for a transducer mounted right next to the head. However, the distance between the head and the transducers will enable reflections to create constructive or destructive interference.

Another evident problem is that phase rotation due to diffraction around the head will increase the travel distance. Thus for angles beyond line-of-sight, an increasing travel time must be expected. This will violate the calculation formula presented in equation 2.4. Figure 2.7 illustrates the issue. The figure shows that the travel length between the receivers consists of two parts, a straight lined trajectory(x_1), and a curved trajectory(x_2) that follows the shape of the head to the transducer. This will have to be compensated for when calculating the angle based on the time difference between ears. Further discussion of the subject will be provided in Chapter 4.1.3.

Figure 2.8 shows the selected locations of transducers. The transmitter is suggested located at the top of the head, such that sound is free to propagate in any direction. As mentioned earlier, the receivers will be affected by shadowing effects since the diffraction around the head will be less prominent for high frequencies, leaving the width where both receivers are are able to receive the transmitted signal narrower. The width of the angle where both receivers will have line-of-sight to the transmitting node will be approximately $\pm 33\,^\circ$ both in the front and the back of the model, depending on the size of the head of the user.

2.3.2 Estimation of distance

Distance estimation is a well-researched subject for ultrasonic positioning. As mentioned in Section 1.3, many systems use an RF signal to synchronize transceivers before an ultrasonic pulse is sent. Thus the time-of-flight (TOF) may be calculated

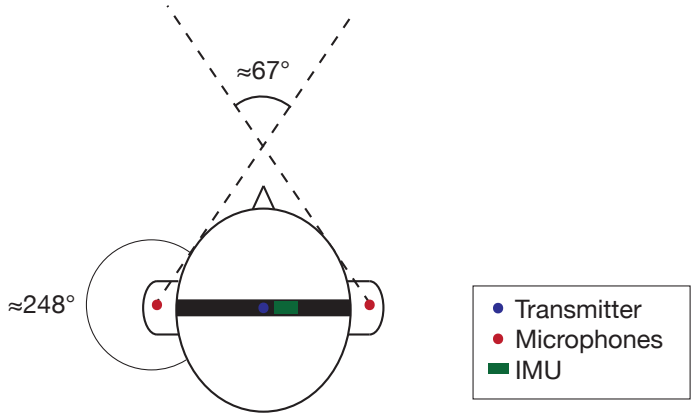


Figure 2.8: Head model illustrating the transducer locations.

as the difference between the RF synchronization signal and ultrasonic pulse. In the system proposed in this work, the positioning is constrained to the ultrasonic domain only. Thus, to estimate distance, a signal needs to be transmitted and received at the node of interest before the response pulse may be used to calculate the distance.

Consider a pulse sent from Node 1 at $t = 0$. The pulse will be received at Node 2 at t_d, the travel time of the pulse. The processing of the signal in Node 2 introduces a processing delay t_p before a new pulse is sent. Thus, the TOF at the pulse at Node 1 is

$$t_{TOF} = 2 \cdot t_d + t_p \tag{2.5}$$

The distance x travelled by a sound wave for a time t is given by

$$x = c_s \cdot t \tag{2.6}$$

Thus, the travel distance may be calculated as

$$x = \frac{1}{2} \left(t_{TOF} - t_p \right) \cdot c_s \tag{2.7}$$

where c_s is a function of temperature as stated in equation 2.2. For two transducer at an angle θ to the source node, the TOF will be different for the transducers. To estimate the distance at the centre of the receiving node, a simple mean may be taken, such that the average TOF is given as

$$\overline{t_{TOF}} = \frac{t_{TOF,1} + t_{TOF,2}}{2} \tag{2.8}$$

This equation may be imprecise for angles outside line-of-sight as the travel distance will increase for one of the transducers (head shadow is discussed in the previous section). However, as the distance accuracy is not a primary concern, the simple formula in equation 2.8 will suffice for the proposed system.

2.3.3 Pulse time resolution

As the angle calculation obviously requires the highest accuracy regarding the time resolution of the pulse, the constraints following from the transducer locations and spacing will define the minimum time resolution requirement.

Consider the resolution constraint of 1° for on-axis positioning as stated in Section 2.1.1. Equation 2.4 defines the angle related to the time difference. Solving the equation for Δt, the given time difference for an angle is given by

$$\Delta t = \frac{d}{c_s} \left(\sin\left(\theta + \Delta\theta\right) - \sin\left(\theta\right) \right) \tag{2.9}$$

Inserting $\theta = 0°$, $\Delta\theta = 1°$ and $d = 0.22$, the time difference becomes $\Delta t = 11.3\mu$s.

For the 90° incidence angle, inserting $\theta = 80°$ and $\Delta\theta = 10°$ yields $\Delta t = 9.8\mu$s.

Thus, the minimum time resolution requirement for the system is

$$T_{res} = 9.8\mu\text{s} \tag{2.10}$$

2.3.4 Head rotation correction

As mentioned in Section 2.1.6, the acceptable delay is as low as 60ms for the HRTF update. Section 2.1.2 emphasizes the subject's own head rotation as the cause of the highest angular speed. It is ambitious to ensure update rates for an ultrasonic positioning system this high. Instead of aiming at a very fast ultrasonic refresh rate, a local *Inertial Measurement Unit* (IMU) may be exploited to correct the orientation of the subjects head orientation.

IMU's may be based on several sensors, providing orientation parameters. If integrated into the subjects head protectors, the given nodes map of surrounding nodes may be rotated independently of the ultrasonic position updates. IMU's does not use global positioning such as GPS as a reference, but rather sensors like gyroscope, accelerometer and magnetometer. Thus, drift may occur over time, but periodic synchronization with the ultrasonic system will provide updated positions for each refresh, and the drift will be insignificant. Using smart filtering, like the Kalman filter, may provide smooth transitions during update of the ultrasonic positions. A kinetic model based on probability may also be incorporated into the system, ensuring that unexpected measured results may be discarded.

2.3.5 Transducer characteristics

The directional characteristics of both transmitters and microphones will affect the performance and functionality of the system. Near-omnidirectional ultrasonic transducers are available on the market, as used in the *Dolphin* system [5]. These transducers do however suffer to low efficiency and are relatively large and bulky. Omnidirectional characteristics would render possible a wider angle of operation for the same signal-to-noise ratio requirement, and also reduce the influence of reflections if the head is directed towards an obstacle or wall.

Directional transmitters and receivers are easily accessible, low cost and easily implemented. If visual contact during conversation is a premise, directive transducers may enhance the performance of the system. The fact that the transmitters and the receiver are mounted on the head of the user allows the user to point his attention to communicating nodes for updating of position and angle. However, for the benefit of a general system as intended in this work, directivity is not a desired property.

A solution to the directionality issue for the available low-cost transducers are transmitter arrays. Arrays must indeed be calculated carefully to obtain a low interference system.

Bandwidth and frequency are as well a concern that limits the performance of a system. Most transducers available for transmission are resonating piezo crystals with a diaphragm. The inherent properties of piezo crystals imply that the bandwidth is limited. The transducers are designed for a specific resonance frequency.

2.3.6 Ultrasonic safety

It does not seem to be a unified consensus regarding the safety of airborne ultrasound [15], but guidelines regarding exposure levels exist. The International Radiation Protection Association[16], recommends 110dB SPL as a restriction for exposure in 1/3 octave bands centered around 40kHz, given continuous exposure during an 8-hour workday. For public exposure, the level is lowered to 100dB SPL. Levels exceeding these limits are allowed if the users wear hearing protectors.

For the proposed system the ear protectors are an integrated part of the system, which reduces the level of exposed ultrasound. However, the transmitter location the top of the head is close to the user's ears, such that there is a chance of hazardous exposure when the equipment is put into the head.

As an example, consider a typical ultrasound transducer with 120dB SPL at 0.3m. At 10cm, the pressure is 130dB SPL. This is the distance between the users ears and the transducer when a top-mounted transducer is used. The pressure is higher than recommendations allows, but considering that normal use requires the use of applied ear protectors, attenuation of 30dB should be expected. For ultrasonic sound, attenuation may prove to be even higher. Thus, the exposure level during operation is below the recommendations. The directional characteristics of the transducer will also influence the exposed level to the users ear.

To solve the issue of exposure during dressing of the system, a mechanism ensuring that the transmitter is idle before properly mounted on the user may suffice. The exposure applied to others, considering the same transducer, will be in the magnitude of 110dB SPL at 1 meter, which is on the limit of recommended exposure for persons not wearing ear protectors. However it should be assumed that for the system to be applied, the level of noise is at such a magnitude that ear protector should be applied.

2.4 Signal transmission technology

To calculate the distance between two nodes, the transmission and reception of a pulse is necessary. The pulse must be transmitted at a given time, detected at the receiver node, and responded to after a constant delay. In addition to the timing pulse, the nodes need a way of communicating their identity. For a multi-node system, the purpose and address of the pulse also needs to be communicated, and as such, a multi-user channel access method or protocol needs to be employed.

A wide range of methods exists to achieve this goal. Digital systems are by far the most adopted and developed systems, as they support a more versatile range of applications and enables complex signal processing. For the prototype, digital processing is the only option available due to time constraints.

To design the signal transmission system, digital communication theory for wireless RF communication is assessed. These systems operate at frequencies much higher than ultrasound, and as well at higher bandwidths and with significantly higher transmission speeds. The communication channels are thus quite different, and various issues may apply to the systems, especially regarding channel modeling. However, the general theory of wireless RF communication may be applied to an ultrasonic system, and the extensive research is beneficial for the ultrasonic application.

The fundamental issues that restrict the ultrasonic positioning system is obviously the channel complexity and the time precision of a pulse. Contrary to a typical wireless communication system, data throughput is not a concern.

2.4.1 Signal design

The ultrasonic positioning systems need to transmit and receive pulses with sufficient precision in time. From signal theory, we know that an infinitely short pulse requires an infinite bandwidth. From a practical point of view, neither of these are possible. Instead, different techniques may be applied to create a pulse with a finite bandwidth and a finite time precision. Methods also exist to gain additional resolution using pulse compression. Pulse compression may be performed on linear or non-linear chirps, or even time-frequency coded waveforms like pseudo-noise codes.

Considering that each node needs to be identified when sending a pulse, a well-known technique that codes an identity based on different sets of spreading pseudo-noise codes may be used. The technique is called *Direct Sequence Spread Spectrum* (DSSS). Each user is assigned a unique spreading code for signal modulation, and the receiver comprises a correlator bank for each user. The technique is developed to spread a message signal onto a wider frequency band than required for the message. Some important advantages of this method are

- the signal may be hidden below the noise floor as the system comprise a processing gain
- each unique code identifies the user, and the signal appears as noise to receivers without a copy of the code
- multi-path challenges may be mitigated or even exploited
- wide bandwidth increases time resolution and is less susceptible to transients
- multiple users may share the same bandwidth (this will be discussed in the next section)

These advantages make the spread spectrum technique useful for a localization system. Several examples utilises this technology. DSSS is for instance used in GPS[17]. Many local positioning systems (LPS) operating in the ultrasonic domain also utilize DSSS, for example, the Dolphin system [5], the system by Urena et. Al [18], and the 3D-LOCUS system [19]. Research has also been conducted on the frequency hopping technique (FHSS) in [6], which provides the same advantages in a slightly different scheme.

The implementations vary significantly, with some using RF in addition to ultrasound such as [5]. Most of the systems are designed for permanent installation in a room to provide very accurate absolute positions. As the DSSS method is well-proven and well described in the literature, it is selected as the channel coding scheme in this work. A detailed description of the theory behind DSSS may be found in for instance [20] or [21].

2.4.2 Channel access methods

Multi-user access is an important property of the system discussed in this work. A communication system has mainly two resources, time and frequency/bandwidth. Various methods for multi-user channel access exploits and shares the resources in different ways, and the most important schemes for wireless communication are described below:

- FDMA (*Frequency Division Multiple Access.*)
  Each user is allocated a frequency sub-band for its use in the overall system bandwidth. FDMA is used extensively in satellite communication and continuous analog systems. For ultrasonic applications, the available bandwidth is often limited, increasing the need for advanced estimation of the pulse acquisition time.

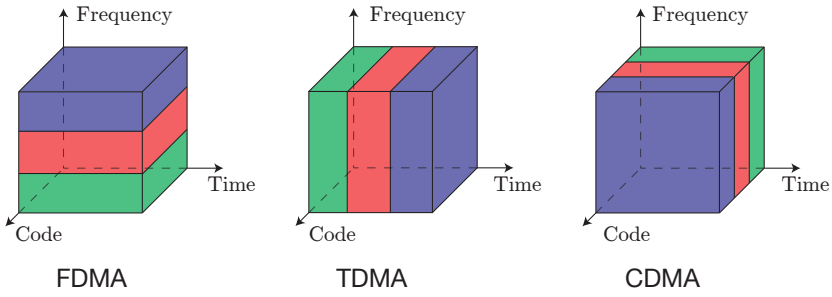


Figure 2.9: Multiple access schemes.

- TDMA(*Time Division Multiple access.*)
  Users are separated by time slots. Only a single user may transmit during a slot. The user is allocated the full channel bandwidth when transmitting. This requires collision detection or a strict slot assignment by a master controller. For ultrasonic applications, time slots will be affected by the distance between nodes. As well, the particular application in this work is desired to be without a master node.

- CDMA(*Code Division Multiple access.*)
  This scheme is in principle a hybrid of FDMA and TDMA. All users share the same time and bandwidth while a spread-spectrum technique such as frequency hopping or direct sequence (DSSS) spreads the signal inside the frequency band.

The different multi-user schemes are illustrated in Figure 2.9. A more thorough description of multi-user access modes may be found in [21, Chapter 8] or [20, Chapter 14].

In radio communications, multiple-access schemes are often related to a base-station-to-user scheme, and allocation of time slots or frequency bands are controlled by the base station. In this work, the system is constrained to not include such features, as each user must be able to communicate with all the other users. Thus, a CDMA system provides a simple approach regarding logic. As a spreading code channel is already selected, it is easily extended to *DS-CDMA*.

Several of the previously researched systems uses DS-CDMA to support multi-user or multi-beacon transmission. This applies to the systems described in [5], [18], [22] and [6]. It is worth noticing that the latter paper describes Frequency-Hopping CDMA, and achieves better multi-path rejection and precision than any of the other positioning systems.

2.4.3 Spreading code

The DSSS principle makes use of a spreading code to increase the bandwidth of a transmitted symbol and obtain the properties discussed in Section 2.4.1. To achieve a maximum spreading of the signal, a *pseudo-random* binary sequence (PN-sequence) may be applied. A PN-sequence has many similarities with noise, but it is periodic and deterministic, such that auto-correlation may be exploited.

A range of codes is found useful for the purpose. Codes include *Maximal length sequences*[20], *Gold codes*[23], *Walsh-Hadamard codes*[20],*Kasami sequences*[24], *Golay pairs*[19] and *Loosy Synchronous codes*[25]. The application of the three latter codes in an ultrasonic positioning system have been analysed in [22]. The different codes have unique properties, and the selection of code depends on the desired qualities, such as cross-correlation properties, auto-correlation properties, inter-symbol-interference (ISI) rejection and available codes per code length. Indeed, also processing requirements and ease of synchronization at the receiver end affects the choice of code.

The most important properties for the design in this work is excellent auto-correlation properties and low cross-correlation properties. As the number of nodes in the system is relatively low, the available codes per code length is not a concern if typical code lengths of 127 to 1022 is selected.

As the main purpose of this work is not to investigate the performance of different PN-sequences, little effort has been put into the choice of code. The Gold code is well documented for use in positioning systems such as in [5], and communication theory literature [21].As well, Gold codes are used in the GPS system[17]. The number of available codes is equal to the code length, and thus ensures flexibility for length selection. Also, Gold code sequences are easily obtained in MATLAB[1]. The Gold code is therefore selected for implementation. However, for a future product development, Kasami sequences or other alternatives may provide better application-specific performance.

The code detection method for DSSS is cross-correlation (CC) between the received code and a corresponding locally generated code. Cross-correlation and its use are assumed known to the reader, but a thorough description of cross-correlation and its properties may be found in for instance [21].

The mathematical definition of the discrete cross-correlation is given as

$$R_{fg}[n] \stackrel{def}{=} \sum_{m=-\infty}^{m=\infty} f^*[m]\ g[m+n] \tag{2.11}$$

for a sequence f correlated with another sequence g, where f^* denotes the complex conjugate of f. In the frequency domain, a convolution in time domain becomes a multiplication. Thus, the cross-correlation may be described as

[1] Matlab will be used for the implementation

$$R_{fg}[n] = \mathrm{DFT}^{-1}\Big\{F[k]G[k]\Big\} \tag{2.12}$$

where

$$F[k] = \mathrm{DFT}\Big\{f[n]\Big\}$$

and

$$G[k] = \mathrm{DFT}\Big\{f[n]\Big\}$$

where DFT is the Discrete Fourier Transform.

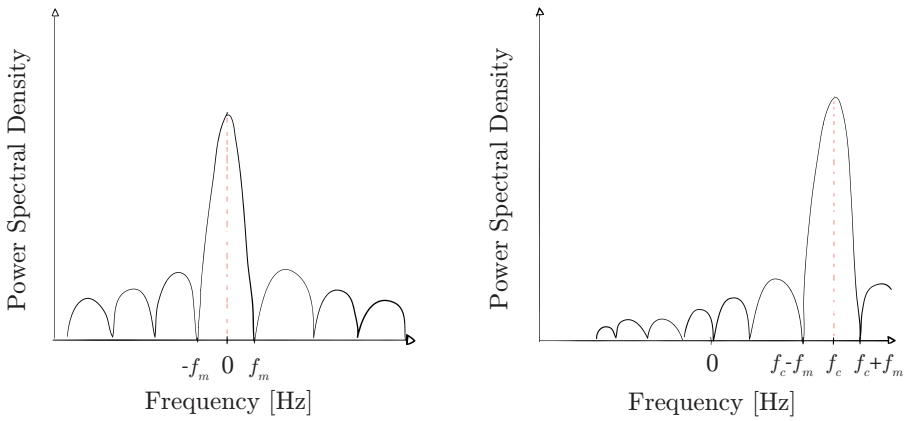
For long cross-correlation input signals, the frequency-domain representation in equation 2.12 may be more computationally efficient than the time domain representation in equation 2.11 due to the efficiency of the Fast Fourier Transform (FFT). For an in-depth review of correlation properties see for instance [21] or the more application-specific note [26], discussing cross-correlation for radar applications.

Considering equations 2.11 and 2.12: If f is a M-length sequence and g is an N-length sequence, the length of the cross-correlation $L = M + N - 1$. In the purpose of calculating the offset between an incoming signal f with a local copy g, zero-padding g to the length of the incoming signal, M, calculation of the offset is easy. The offset between two signals may be found from the cross-correlation as the peak correlation coefficient above a certain detection threshold, and is denoted the *lag*. A zero lag in the cross-correlation $R_{fg}[n]$ is represented by the index $n = M$.

For a communication system, the PN-sequence is used to spread a modulated data signal. In the case of the positioning system in this work, only the code itself is of interest. Thus, the PN-sequence may be modulated directly onto a carrier wave using binary phase-shift keying (BPSK). When a spreading code is used in a DSSS system, the code sequence is denoted a *chip sequence*. As the chip sequence is a binary pulse sequence in the time domain, the resulting frequency domain representation is a sinc-function. The first zero point in the frequency domain sinc function will correspond to the chip rate, denoted f_m.

2.4.4 Transceiver fundamentals

This section will further describe the application of the selected system strategy in a transceiver. Three main system components will be assessed; the transmitter part, the receiver part and the system logic. For further detail, please consider Chapter 3.



(a) Baseband signal (b) Passband signal

Figure 2.10: Illustration of the up-mix process.

2.4.4.1 Transmitter

A pure PN-sequence will, as mentioned earlier, have spectral content as a sinc function in the frequency domain. The sinc function will be centered around DC and contain both positive and negative frequencies, denoted a *baseband* signal. To transmit the signal over an ultrasonic channel, the spectral content must be modulated to an ultrasonic frequency. The process is called *up-mixing*. An illustration of the signal in the frequency domain is provided in Figure 2.10. The baseband signal spectrum is shown in 2.10a, and the corresponding up-mixed passband signal is shown in 2.10b.

The PN-sequence $p[n]$ with values [0,1]is transposed to a binary phase signal by simple algebra such that a chip-sequence for binary phase modulation $m[n]$ with values [-1,1] is created:

$$m[n] = 2 \cdot p[n] - 1 \tag{2.13}$$

A carrier is created at the desired transmission centre frequency as

$$c[n] = \cos\left(2\pi f_c \frac{n}{T_s}\right) \tag{2.14}$$

where f_c is the carrier frequency and $T_s = \frac{1}{f_s}$ is the system sample period.

Thus, the sample rates of the carrier frequency and the chip signal is not equal. To up-mix the signal, a simple multiplication of carrier and phase signal is performed. However, the phase signal first needs to be up-sampled to the system sample rate f_s by a factor

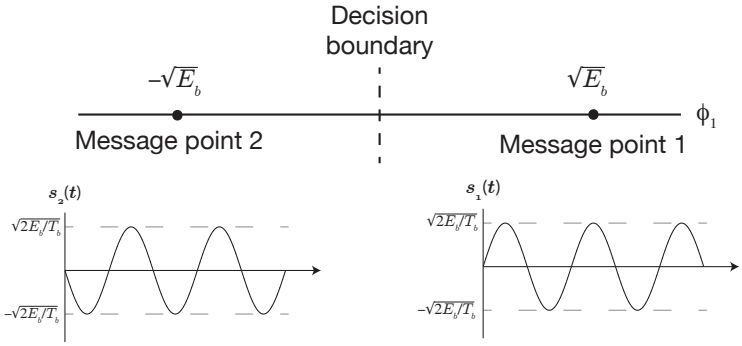


Figure 2.11: Signal space diagram for coherent BPSK system.

$$P = \frac{f_s}{f_m} \tag{2.15}$$

Adding $P-1$ samples between each sample in $m[n]$ and setting all the consequent added zeros to the corresponding first sample value in m provides an up-sampled phase signal $\dot{m}[n]$ (note the dot, indicating an up-sampled signal). Thus, up-mixing by multiplication may be performed as

$$s[n] = \dot{m}[n]\ c[n]. \tag{2.16}$$

Up-sampling in the digital domain will require oversampling and additional complication if P is not an integer value. The sample-rate to chip-rate $\frac{f_s}{f_m}$ is therefore usually selected to yield an integer value for P.

An illustration of signal-space diagram for BPSK modulation is given in Figure 2.11, along with the resulting waveform. E_b is the energy per chip. More theory regarding BPSK modulation may be found in [21, Chapter 6]. An excerpt of a Gold PN-sequence $p[n]$ is presented in Figure 2.12. The corresponding up-sampled and transposed chip-sequence $\dot{m}[n]$ is illustrated in Figure 2.13. The time-domain waveform resulting from up-mixing is illustrated in Figure 2.14 where the transition points are marked for reference, and the corresponding frequency domain representation is plotted as the power spectral density in Figure 2.15. The frequency domain representation is plotted with both negative and positive frequencies.

2.4.4.2 Receiver

The aim of the receiver is to correlate the incoming signal with a local copy for each of the PN-sequences to extract the code identity and the offset of the incoming

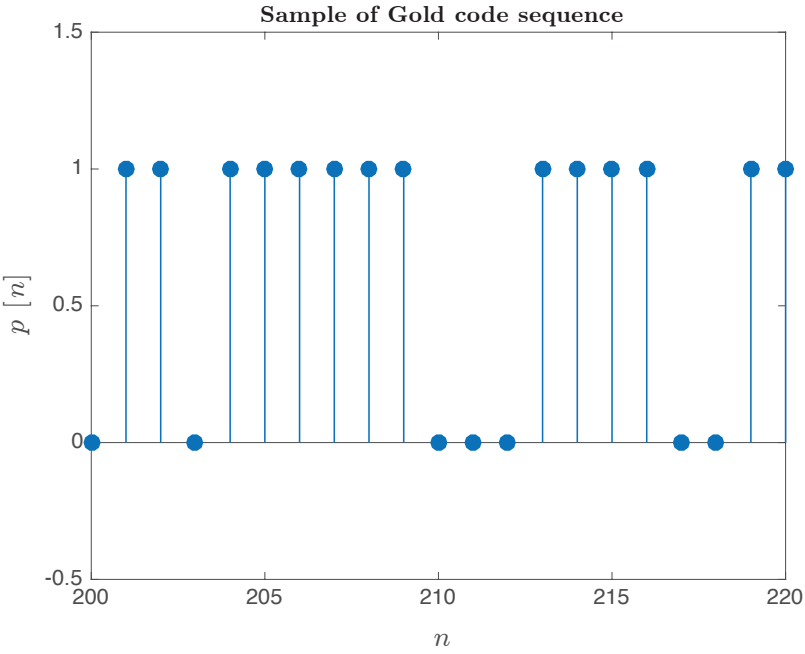


Figure 2.12: Sample of Gold code PN-sequence $p[n]$, 21 bits extract.

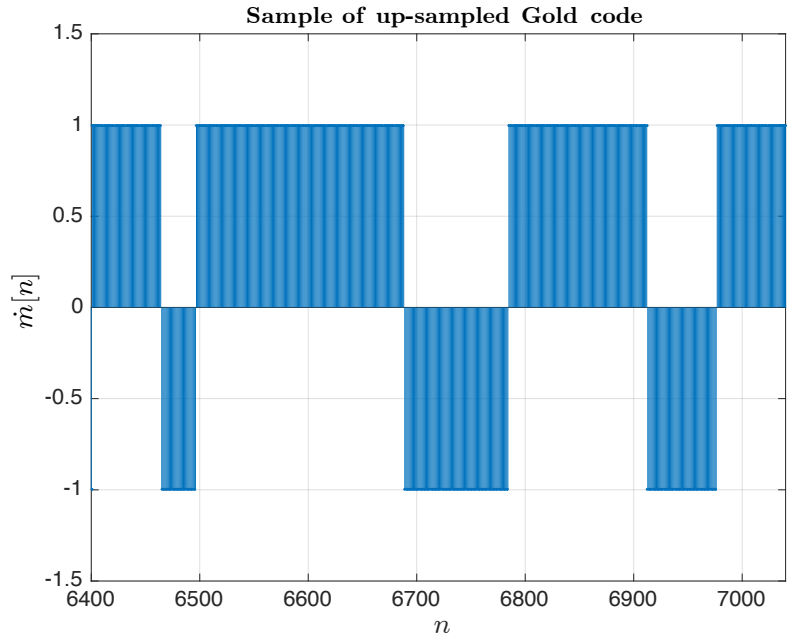


Figure 2.13: Sample of up-sampled chip-sequence $\dot{m}[n]$

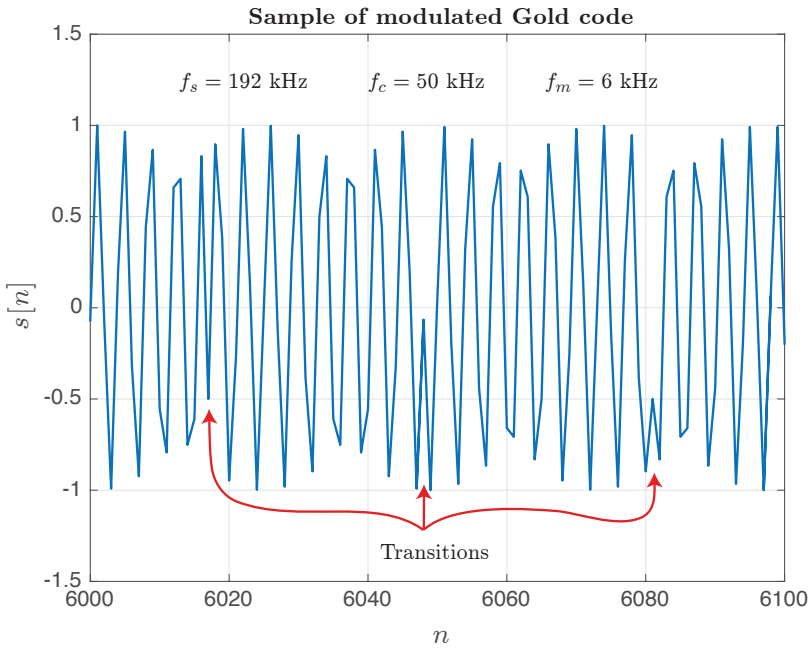


Figure 2.14: Excerpt of the modulated output signal $s[n]$

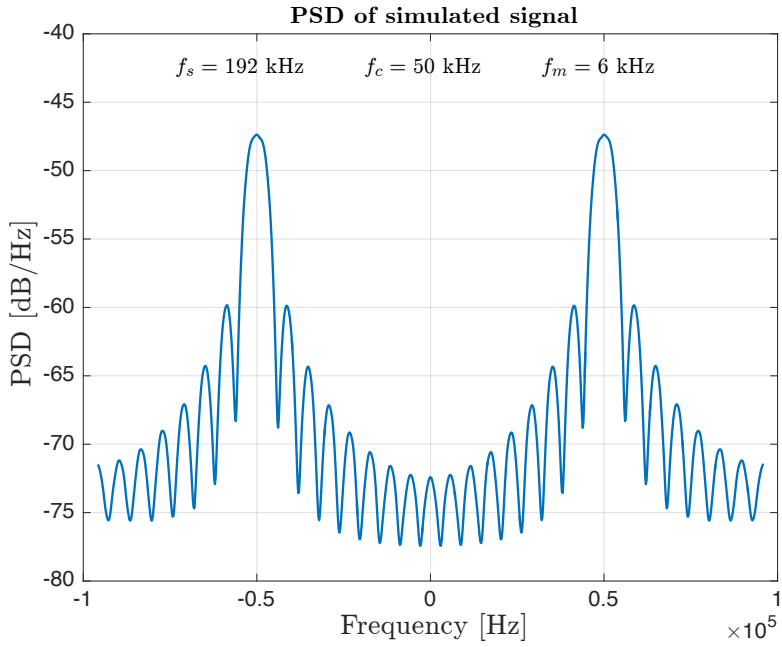


Figure 2.15: PSD of the modulated output signal $s[n]$

signal. The receiver may be divided into two parts; the demodulator and the correlator.

Demodulation
To save computational power and remove excess frequency spectrum, the signal is first down-mixed to baseband and low-pass filtered. Down-mixing is the reverse process of up-mixing, also performed by a local carrier multiplication. However, the phase of the incoming signal $r[n]$ is unknown, and two different strategies may be used to prevent a phase mismatch causing loss of signal:

a) carrier synchronization

b) complex demodulation

Solution a) requires that the phase of the carrier in $r[n]$ is calculated, thus enabling the local carrier $c[n]$ to be generated in-phase with the incoming signal. The principle of carrier synchronization is intuitive and is mostly used in traditional analogue systems using phase-locked loops (PLL). A typical method is the *Costas loop* [27].

Solution b) is based on an unsynchronized multiplication with two 90° offset carrier waveforms, denoted I (in-phase) and Q (quadrature) respectively. Thus, the absolute value of the signal output after multiplication will always be constant, regardless of the carrier phase. The output is a complex signal, and further processing is performed in the complex domain, including synchronization.

After down-mixing the signal, spectral images of signal of interest will exist at frequencies around $f = kf_c$ for integer values of $k = 1, 2, ..\infty$. To reject these images, a low-pass filter is applied to the down-mixed signal. This will also ensure that all signal energy outside the spectrum of interest is reduced. For the real signal r_{dm} obtained by method a), the filter is a real LP filter. For the complex signal $\boldsymbol{r}_{dm}$ obtained by method b), two separate filters for the I and Q paths are used. The output of the filters are denoted r_{LP} and $\boldsymbol{r}_{LP}$ respectively.

To illustrate the concept, the power spectral density of r_{dm} is plotted in Figure 2.16. A corresponding low-pass filtered signal is illustrated in Figure 2.17. Note that no sidelobes are present due to a very sharp filter. The filter ensures that decimation may be applied instead of an integrate-and-dump detection strategy.

The theory and practical implementation of carrier synchronization, as well as symbol synchronization, is the most complex part of the demodulation process. A lot of research has been conducted on the subject, and the literature is often directed at specific implementations. However, most research is focused on a base station-to-user scheme, where continuous synchronization may be performed on the incoming signal. Burst transmissions as required for a positioning system is to a less degree covered in the literature. An in-depth discussion of the techniques is out of the scope of this work, but references for further studies are included below.

In [20, Section 13.2.3] a short introduction to synchronization for DSSS systems may be found. The acquisition, tracking and demodulation of GPS signals are quite similar to the practical application in this work, and is thoroughly discussed in [17, Chapter 5]. A practical summary of different synchronization methods may be found in the lecture notes in [28], with a focus on digital methods for

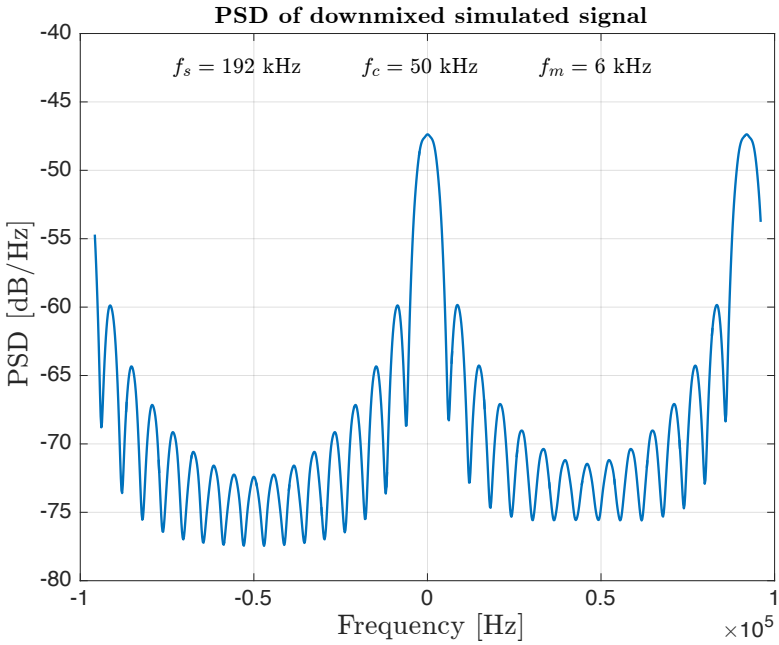


Figure 2.16: PSD of the down-mixed received signal $r_{dm}[n]$

synchronization. For complex demodulation, a useful description of the principles and annotation is found in [29].

Although the carrier frequency may be assumed constant and known for a general system, different channel effects apply. For instance, a moving node will introduce Doppler shifts of the frequency. Indeed, for low-speed channels such as ultrasound, the frequency shift is a major concern even at a low-speed movement. In [30], the effect of a node movement in a positioning system is studied. In this work, the Doppler shift issue is acknowledged as a limitation, but no further effort is dedicated to this challenge. Further work may implement tracking mechanisms to mitigate the issue. Frequency estimation for burst transmission is covered in for instance [31].

Another significant channel issue is the *near-far effect*. For two nodes transmitting to a third node at the same time, the distance to the third node will affect the signal level at the receiver. As each chip sequence appears as noise at the correlator for another chip sequence, the SNR threshold will determine if the signal from the further node may be received. In wireless RF systems, the signal output level is often adjusted to mitigate the issue. For an ultrasonic decentralised system, this will be a complex or impossible way to mitigate the issue. Instead, increasing the processing gain of the system by longer codes may require lower SNR for detection.

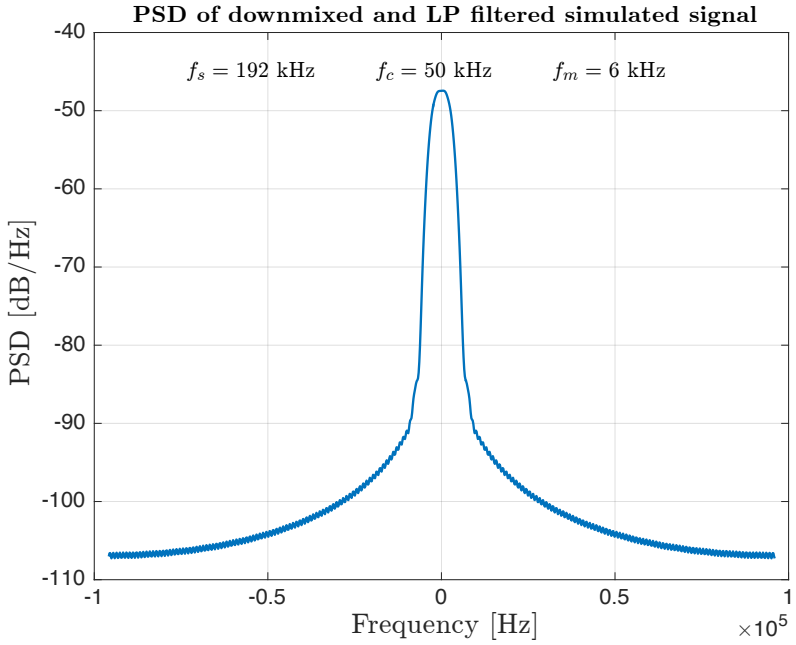


Figure 2.17: PSD of LP-filtered down-mixed received signal $r_{LP}[n]$

Correlation

The purpose of the correlation step is to obtain the chip sequence identity of the incoming signal along with the pulse arrival time. This way, the TDOA and TOA may be calculated, and angle and distance may be estimated as discussed in Section 2.3.

After demodulation, the received signal consists of spectral content limited to the chip-rate f_m. Cross-correlation can be performed directly on the phase signal r_{LP} using an up-sampled local chip sequence at f_s. This requires chip sequence lengths $L_{chip} \cdot P$, with P given by equation 2.15. An alternative approach, saving considerable computational power, is to synchronize the chips and integrate over the message symbol to obtain a chip-rate sequence. A simple decimation may suffice given the received signal is low-pass filtered and the decimation is synchronized sufficiently. However, chip synchronization is, as discussed above, a complex subject which requires advanced signal processing.

A cross-correlation R_{fg} of a chip sequence $f[n]$ with 30 zeros padded before the sequence begins, and the same chip sequence without padding $g[n]$, is plotted in Figure 2.18. Both sequences contain a Gold code with length $L_{chip} = 511$ and is padded to a total length $N = 2L_{chip}$. The plot shows how the cross-correlation peak is positioned exactly at 30 lags.

The cross-correlation and lag calculation may be performed as described in 2.4.3 and equation 2.11 and 2.12. The correlation coefficient serves as an indication of

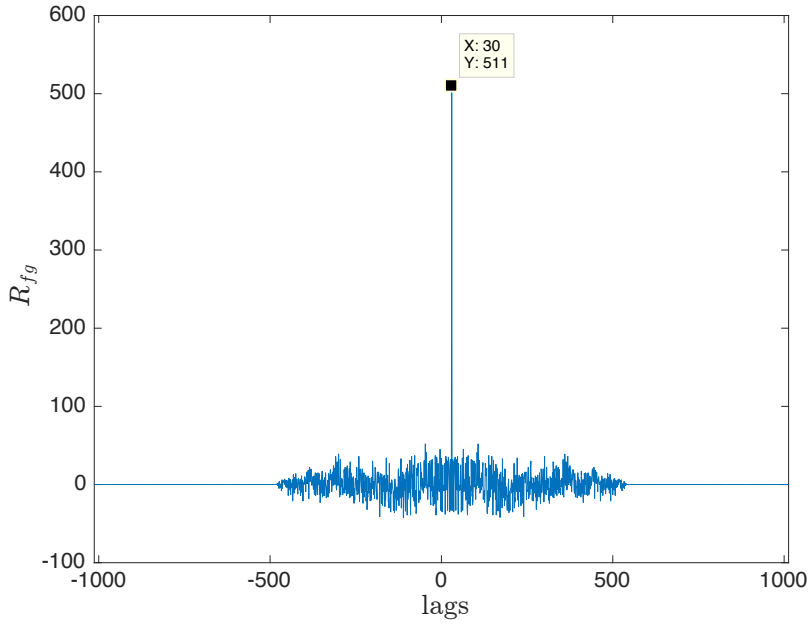


Figure 2.18: Cross correlation of two equal Gold sequences f and g, f lagging g by 30 samples.

how strong the correlation is, and thus is used to assess whether a sequence is detected or not.

2.4.5 Signal transmission logic

As mentioned earlier in this chapter, the chip sequences exploited in the DSSS system may be used to identify different nodes, and as well provide specific commands. The positioning system depends on a call-response method. Thus, to obtain the positions of surrounding nodes, a node has to issue a request and receive an answer. Several schemes may be used to obtain this functionality, but developing a smart scheme may increase the efficiency of the system considering the total number of transmissions and update rate. Some important aspects are listed below:

a) For a system without a centralised management system, each node has to assign itself an identification number.

b) The refresh rate may be set such that transmissions are prevented if a given node's position has recently been updated.

c) If a call-response has been performed, only a single additional transmission may provide position information for the other node as well.

Consider a one-to-one signalling scheme with nodes N_k where $k = [1, K]$. Each

node has two commands (represented by unique chip sequences) to transmit, the Request Q_n and the response S_n. Each node will acquire a position by sending a request to another node. Let's say N_1 needs the position of N_3, the following sequence is necessary: $N_1 \rightarrow Q_3$; $N_3 \rightarrow S_1$, where $\rightarrow$ denotes a transmission. Thus, a single position needs two transmissions. For all nodes to estimate the position once, each node will have to ask for three positions - thus, the number of transmissions will be

$$N_{tr} = K \cdot (K-1) \cdot 2 = 2\left(K^2 - K\right) \tag{2.17}$$

For four nodes, the resulting number of transmissions will be $N_{tr} = 24$, and the number of commands needed for each node would be two. This scheme is not efficient, but could, if desired, allow for a TDMA multiple access scheme if some kind of RF time synchronization and identity negotiation is possible.

Instead, aspect c) above could be exploited. If each node has three commands available, Q_n, S_n and a retransmission command R_n, the position for two nodes can be determined by the following sequence: $N_1 \rightarrow Q_3$; $N_3 \rightarrow S_1$; $N_1 \rightarrow R_1$. With this scheme, all positions will be obtained by

$$N_{tr} = 3\,(K-1) \tag{2.18}$$

transmissions, thus for four nodes $N_{tr} = 9$.

Aspect a) above suggests that the determination of an identity should be an easy operation. If a single broadcast triggers response from all nodes (requires a CDMA scheme since simultaneous transmission will be a consequence), this requirement will be fulfilled. Thus, the following sequence would be necessary for a three-node system: $N_1 \rightarrow Q_1$; $N_2 \rightarrow S_2$; $N_3 \rightarrow S_3$; $N_1 \rightarrow R_2$; $N_1 \rightarrow R_3$. Here, three commands are required by each node. The communication is illustrated in Figure 2.19. To obtain positions for all nodes,

$$N_{tr} = 1 + 2 \cdot (K-1) \tag{2.19}$$

transmissions is required. Thus for four nodes, $N_{tr} = 7$ transmissions are required to position all nodes once. This scheme is denoted *broadcast-response-retransmission* for further reference.

The last aspect that will minimize the number of transmissions is the aspect mentioned in aspect b) above. If all nodes request positions from other nodes at a random delay after a given refresh rate, the number of transmissions becomes higher than proposed in equation 2.19. However, by keeping track of the last refresh of a position, the answer S and retransmission R may be omitted if the position of the requesting node was recently updated. The random delay will also ensure that collisions will be random, thus reducing problems resulting from near-far effects and overlapping transmission.

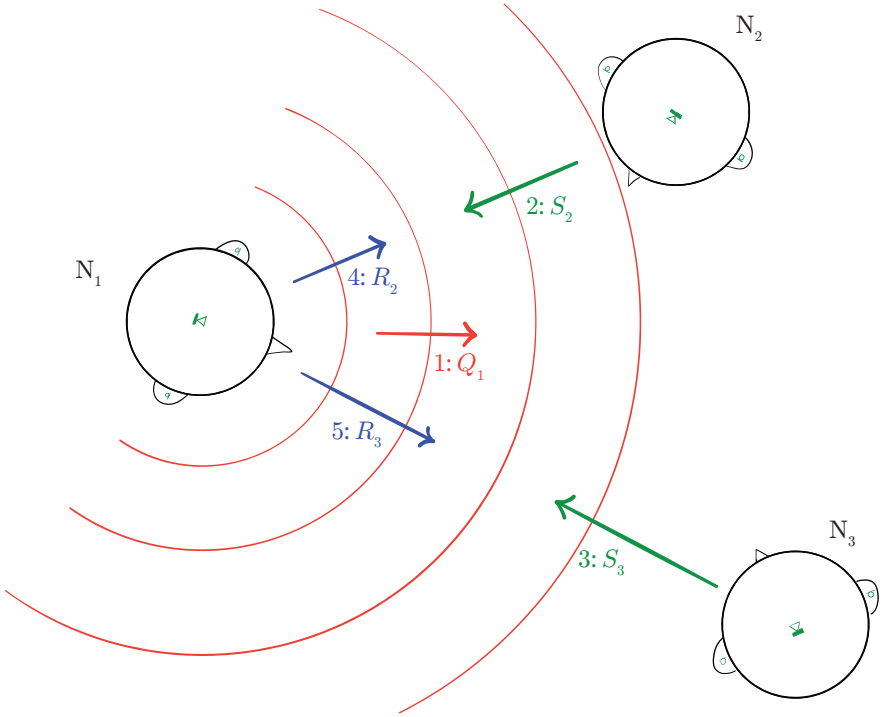


Figure 2.19: Illustration of the broadcast-response-retransmission scheme for three nodes. The command pulses are denoted Q for request, S for response and R for retransmission. The number before each command indicates the order of code transmission start (Node 3 is further away from Node 1 than Node 2).

2.4.6 Link budget

To estimate the range and robustness for the system, a link budget is calculated. The link budget is obtained by quantifying the parameters that will result in gains and losses for each link in the system chain; this includes the transmitter, the transmission medium and receiver. The budget considers a node-to-node transmission.

Link budgets for ultrasonic communications systems are described in for instance [32]. All parameters in the signal chain is included; the source level (SL), the propagation loss (PL), the noise level (NL) as well as the detection level (DT) necessary to achieve a successful decoding of the signal. By adding a term for propagation gain(PG), the benefits of the DSSS system is included. The budget is defined by the following equation:

$$SL + PG - PL - NL > DT \tag{2.20}$$

SL equals source level in dB SPL at referenced distance, usually $x_0 = 1$m. This factor will be restricted by the efficiency of the transmitter and the available power in the device.

PL is the propagation loss, mainly caused by the distance from the transmitter and attenuation caused by air absorption. The air absorption depends on the relative humidity in the air, temperature and pressure of the transporting medium. According to Holm [33] and Figure 2.3, the maximum attenuation due to air absorption for a frequency 40kHz at 20°C is around $\alpha = 1.33$dB/m at a relative humidity 55%.

The intensity of the sound pressure (power of the signal) at the receiving node(I_i) decreases quadratically with distance(x) in relation to the reference distance and Intensity I_0, such that

$$I_i = I_0 \left(\frac{x_0}{x}\right)^2 \tag{2.21}$$

This yields a loss of $20\log_{10}(x)$. The directivity of the transmitter may also be accounted for. For this link budget, only an omnidirectional response is considered. Thus, the total propagation loss is $PL = 20\log_{10}(x) + \alpha x$.

The amount of background noise level (NL) must be estimated. According to Bass and Nolan [34], machinery in the process industry produces substantial background noise in the ultrasound range, most prominent is noise caused by aerodynamic effects. Levels as high as 80dB SPL at 500kHz measured at 1m was reported. In the $20 - 60$kHz range, an average level of 70 - 80dB SPL was reported in an industrial environment measured at a bandwidth of 3kHz. An equivalent spectral density[dB/Hz] can be calculated as $NL = 70 - 10\log_{10}(3000) = 35$dB/Hz.

As may be seen from the calculation above, the total noise level will be affected by the bandwidth of the transmitted signal. A wider bandwidth yields more noise energy at the receiver.

According to Holm et.al[32][35], the detection threshold DT can be assumed to be in the range of 15dB to 20dB for optimal detection theory.

As discussed in Section 2.4.3, spreading the signals will give a benefit in form of a processing gain. The processing gain is defined as

$$PG = 10\log_{10}(L_{chip}) \tag{2.22}$$

Using PG and DT, the signal-to-noise ratio may be estimated as

$$\text{SNR} = DT - PG \tag{2.23}$$

The link budget equation may set the requirements for source level and processing gain:

$$SL + PG - PL - NL > DT \tag{2.24}$$
$$SL + 10\log_{10}(L_{chip}) - 20\log_{10}(x) - \alpha x - 35 - 10\log_{10}(BW) > 20 \tag{2.25}$$
$$SL + 10\log_{10}(L_{chip}) - 10\log_{10}(BW) > 55 + 20\log_{10}(x) + \alpha x \tag{2.26}$$

Thus, for a distance design goal of $x = 50$m as described in Section 2.1.1, the requirement will be

$$SL + 10\log_{10}(L_{chip}) - 10\log_{10}(BW) > 156 \tag{2.27}$$

If instead the ultrasonic background noise is considered to be as low as 20dB SPL in the same spectrum, the resulting requirement will be

$$SL + 10\log_{10}(L_{chip}) - 10\log_{10}(BW) > 96 \tag{2.28}$$

2.5 Design parameters and summary

Based on the discussion of the model and requirements in this chapter, a summary of parameters is presented below, in Table 2.1. These parameters will be used in the implementation under the constraints of the available hardware.

Table 2.1: Summary of design parameters

Parameter	**Symbol**	**Value**	**Section**
Number of mics		2	2.3.1
Mic spacing	d	0.22m	2.3.1
Mic location		4mm above earmuff	2.3.1
Pulse time res.	T_{res}	9.8μs	2.3.3
Range	x_{max}	50m	2.1.1
Signal design		DS-CDMA	2.4.1
PN-sequence type		Gold code	2.4.3
Code scheme		Broadcast-response-retransm.	2.4.5
Noise spectral density	N_0	35dB/Hz	

Chapter 3

System Implementation

This chapter presents and describes the implementation of the positioning system, based on the design choices in Chapter 2. The system will be discussed in terms of the prototype implementation and its limitations. Critical parameter selections based on available hardware will be presented. Theory of operation for implementation-specific components will be discussed along with the chosen methods. This chapter also includes simulation results to support the choice of methods.

3.1 Prototype system architecture

To achieve a system-level real-time prototype during the time available for this work, digital processing in high-level signal processing environment is one of the main constraints. An environment familiar to the authors was crucial. Mathworks' MATLAB was selected for this purpose, as it provides many relevant functions for signal processing and communication systems.

The principal setup of a system node is a computer running MATLAB. The computer is connected to the node transducers via an audio interface. The node hardware is a set of two microphones and one transmission transducer, both mounted on an artificial head. The head and transducers are supported by a microphone stand. The transducer connections for each node is presented in Figure 3.1. In addition, an IMU (Inertial Measurement Unit) input is connected to the computer via wireless network. The IMU is implemented as a MATLAB application running on a smartphone, using the phone's sensors to calculate the orientation. The phone is attached to the artificial head. A picture of the node setup is included in Figure B.7 in the appendix.

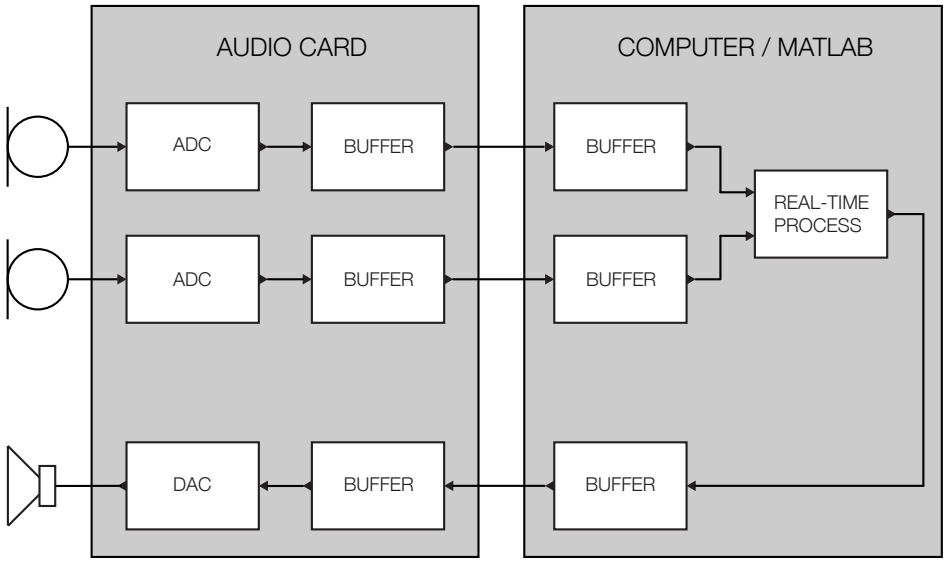


Figure 3.1: Overview of node processing system

3.1.1 System hardware

As the resources available to this work is limited, the hardware is selected upon availability. However, all hardware satisfies requirements for an ultrasonic system. The selection has not compromised the goal, but rather emphasised the constraints imposed by real system components. A complete list of equipment and appurtenant serial numbers are presented in Appendix B, but the hardware and their characteristics will be presented below.

3.1.1.1 Artificial head

The artificial heads used for the prototype was of type Sennheiser MZK-2002. These are mannequin heads widely used for binaural recordings. The shape of the head possess the natural characteristics of a human head, with approximately average dimensions. Verification of the form and size showed that the height from the nasal root to the top of the head was approximately 30mm shorter than an average male head (see section 2.1.4). This implies that the ears are closer to the top of the head than normal, and could affect the measurements by means of a different wave propagation path than for an average head. The model in section 2.3.1 is based on wave propagation around the head as the shortest path, a shorter distance from ears to the top of the head may cause the shortest propagation path to be above the head instead. Thus, erroneous results may be expected for large angles. Different microphone positions are examined in Chapter 4. For the benefit of easy assembling and accurate positioning of the microphones, two microphone stands were designed and built by the engineering workshop.

3.1.1.2 Turntable

To be able to precisely rotate and control the angle of the node a turntable was used. The equipment used was a Norsonic Nor265, which is capable of rotating an axle to a particular absolute angle. The turntable provides control over a standard serial port, and for the prototyping, a control algorithm was implemented using MATLAB. Settings include speed and acceleration, and the turntable is also capable of reporting the current angle.

3.1.1.3 Computers and software

Running a real-time code in MATLAB on a laptop computer introduces certain constraints to the computational load. A thorough calculation of capabilities is out of the scope of this work, but calculation limitations were encountered during the implementation phase and dictated the further path of development.

The specification for the computers used for the prototype is included below for reference:

Computer 1	
Computer type:	Apple Macbook Pro 15" 2013
Processor:	Intel Core i7 2.6 GHz
Memory:	16 GB 1600 Mhz DDR3
Computer 2	
Computer type:	Apple Macbook Pro 13" 2015
Processor:	Intel Core i5 2.7 GHz
Memory:	8 GB 1866 MHz LPDDR3

Both computers was running MATLAB R2015b 64-bit for Mac Intel OS.

3.1.1.4 Audio interface

Two different audio interfaces from the same manufacturer was used. Roland Octa-Capture and Roland Studio-Capture. These modules use USB 2.0 to interface the computers, and may process at sample rates 192kHz, 96kHz, 48kHz and 44.1kHz. The interfaces were verified to have a flat frequency response up to about 60kHz using the 192kHz sample rate configuration.

3.1.1.5 Microphones

Microphones used in the implementation was selected out of availability. Four identical capsules were not available, and the two nodes in the test setup use different types of capsules, though pairwise identical. *Bruel & Kjær* 4939 and 4938 was used. Both types are 1/4" capsules with a specified frequency range of up to 100kHz and 70kHz respectively. The polar pattern was measured for the frequency

range of interest and can be seen in Figure 3.2. As can be seen in the plots, both microphones have an attenuation of ≈ -30dB at 90° incident angle, which is expected to affect the results of the off-axis measurements. The characterization of the polar pattern was performed in the anechoic chamber, using the same ultrasonic transmitter as described below. The audio interfaces described above was used for the measurement, along with MATLAB as the signal processor. Though different in frequency range and sensitivity, the microphones characteristics are considered alike to the application in this work.

The capsules are all amplified by the same type of preamp, the Norsonic Nor1201, which is characterized by a flat frequency response of up to 1MHz. Two NOR336 from the same manufacturer is used as power supplies for the microphones. The 4939 microphones was used for the turning node while the 4938's was utilised for the static node.

3.1.1.6 Transmitter

The transmitter used to generate the ultrasonic sound is an off-the-shelf product from Murata of unknown type. However, after the characterization mentioned below, it was assumed to be a MA40S5[1].

The transmitter is based on piezo-electric technology. Contrary to many off-the-shelf ultrasonic transducers, the chosen transmitters does not require a high-voltage amplifier circuit. The maximum input voltage was manually obtained by measuring the sound pressure output from the transmitter using a signal generator. By varying the voltage, the maximum was determined to be 5V_{p-p} as saturation was measured above this level. The measured sound pressure level at 1m on-axis was 91.5dB SPL. However, the data sheet assumed to cover the transmitter suggest levels of $20 - 60\text{V}_{p-p}$ may be induced for short durations, thus giving up to 114dB at 1m on-axis.

To amplify the transmitters, an NAD 312 integrated stereo amplifier was used. The frequency response was verified to be flat up to around 60kHz at no-load conditions. The data sheet of the amplifier specifies a -3dB maximum deviation up to 70kHz.

The transmitters are characterized by their frequency response and directivity, measured in an anechoic room with the audio interfaces and microphones described above using MATLAB as the signal processor. Figure 3.3 shows a maximum sound pressure output at 40.80kHz. The -3dB roll of appears at 40.05kHz and 41.50kHz, thus a bandwidth of 1.45kHz.

Figure 3.4 shows that the directivity is quite narrow, -3dB appears at around $\pm 27°$. The implication is that the transmitter will reduce the system's signal-to-noise ratios off-axis, and possibly augment reflection-related interference. For on-axis angles, the opposite effect is obtained.

[1] See for instance `http://www.farnell.com/datasheets/484560.pdf`

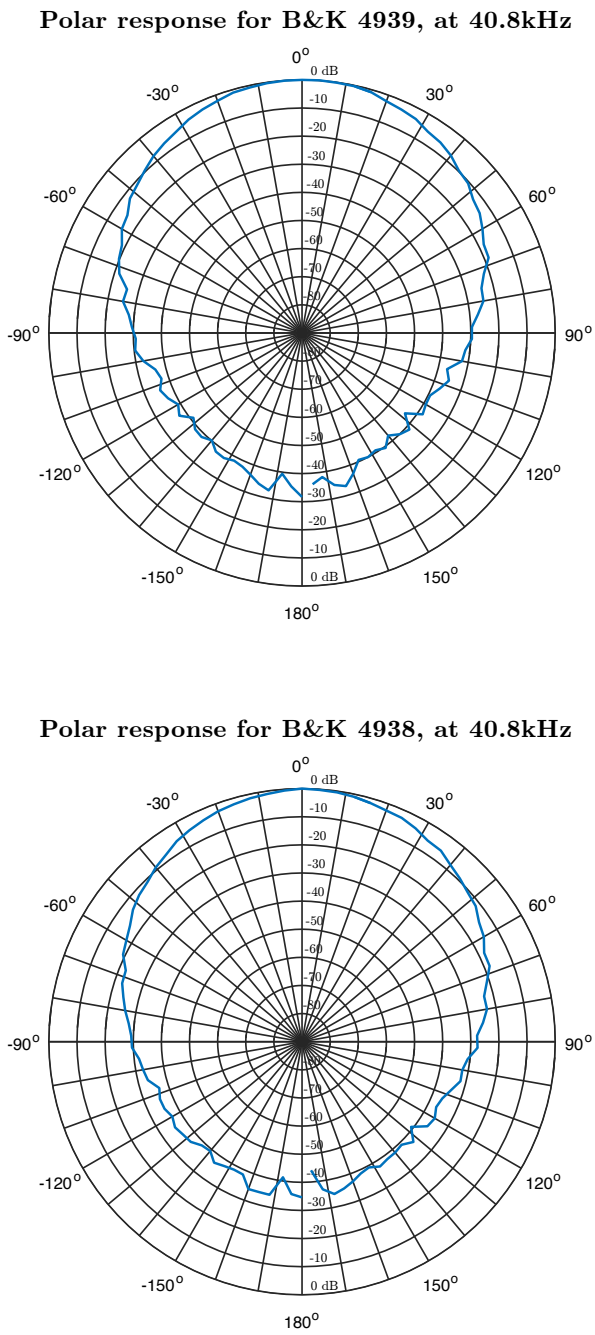


Figure 3.2: Measured directivity for microphones B&K 4939 and B&K 4938 at 40.8kHz, normalized to on-axis response.

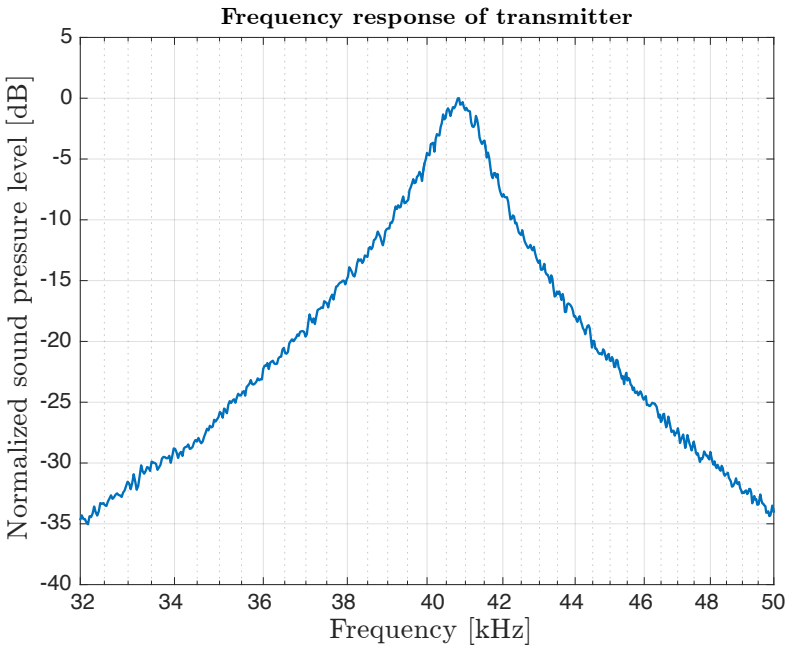


Figure 3.3: Frequency response for the transmitters

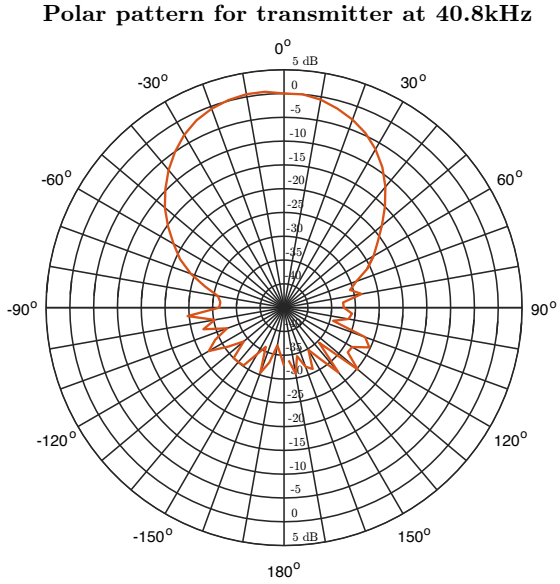


Figure 3.4: Polar response for the transmitters at 40.8kHz

3.1.1.7 IMU signals

The IMU data is implemented as a MATLAB application developed for iOS smartphones. The application is capable of transmitting the orientation sensor output to a desktop computer running a MATLAB instance via a computer network. The network type used in this work is a standard 5GHz WiFi network. The refresh rate is about 2Hz, and the update is delayed by a randomly varying and unpredictable network lag. Synching the angle to the node system for continuous rotation measurements will thus provide unreliable results, and constant rotation angles using the turntable should be used to ensure an accurate reference angle. The sensors are processed on the smartphone's operating system; thus, several sensors contribute to the output (accelerometer, gyroscope, and magnetometer). Only the azimuth angle output is utilized by the system implementation.

3.1.2 Signal processing in MATLAB

The MATLAB code is built up using classes, as an object-oriented programming architecture. The fundamental class running the real-time loop is a modification of a *Mathworks*[2] example code named *Real Time Audio Processor* [36]. In this text, the class will be referred to as `RTAP`. This class holds and administers the incoming and outgoing signal buffers, and provides a graphical user interface (GUI) for setup of audio interfaces and node parameters. In addition, it provides a control routine for a smart-phone based IMU implementation, and a turntable control routine for measurements and analysis of the system. Figure 3.5 shows the relation between the implemented objects.

The `RTAP` class holds a second class implementing the functionality of the communication system, the `Node` class. This class holds all configuration parameters related to a node, such as the node identification number and all the objects required to run the positioning system as a unit. The `Node` also handles the processing buffer (further described in the consecutive section). The following classes are implemented to operate the system, and is held by the `Node` object:

`Modulator`
: Required to generate an output signal based on the chip sequence number.

`Demodulator`
: Required to demodulate the incoming signal from the carrier frequency to a chip-rate signal.

`Correlator`
: Required to correlate the demodulated signal with a local chip sequence set, and find the corresponding time lag.

`Timekeeper`
: Required to keep track of the absolute time and calculate the absolute time of arrival for the received signal pulses.

[2] Mathworks is the company developing and maintaining the MATLAB software.

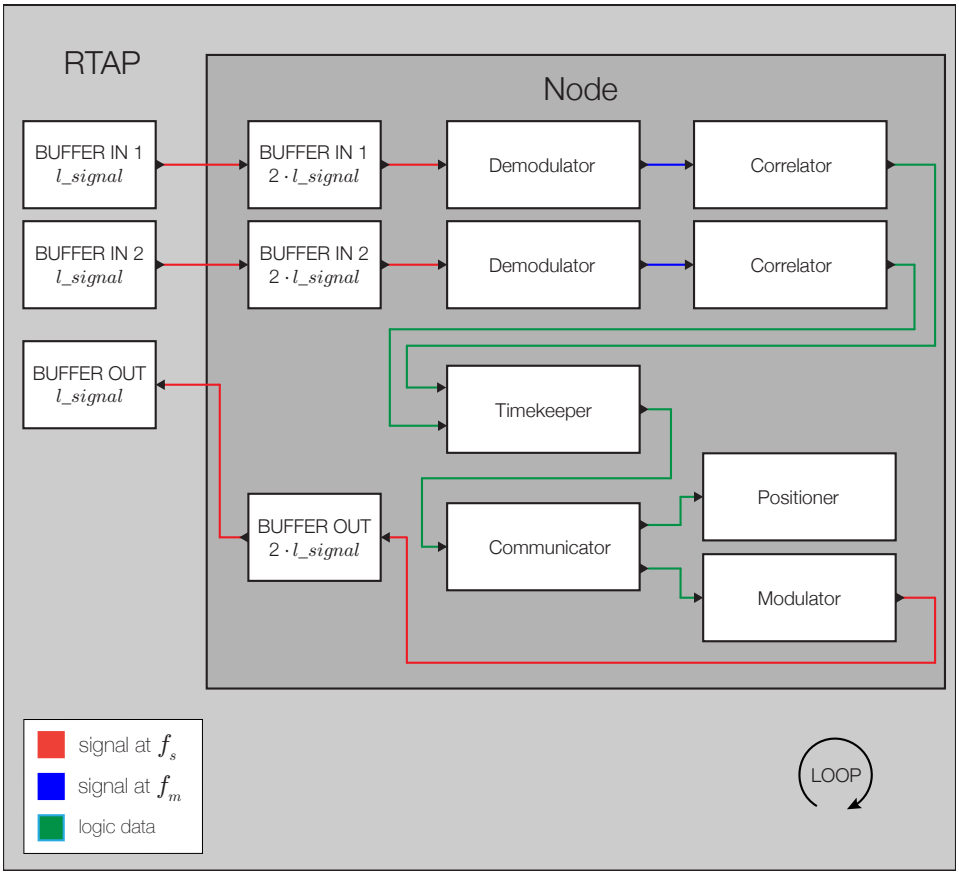


Figure 3.5: Overview of MATLAB class relations

`Communicator`
: Required to calculate distance and angle, and handle the overall system logic.

`Positioner`
: Required to keep track of the IMU signals and the positions acquired from the ultrasonic system.

A block-based approach is chosen due to easier debugging and to share the programming workload between the authors. The modular approach also allows different modules to be tested, such that several implementations may be evaluated. In the final implementation, the Demodulator and Correlator are merged into a single class called Detector.

The classes will be discussed on a conceptual basis in the following sections. A class reference description is included in Appendix A.

It is important to clarify that the prototype implementation may be rather inefficient for an embedded application. For instance, processing the signals in buffers are not necessarily the better option, but in the MATLAB implementation the buffer method must be used. Embedded processing may thus require significant redesign, and as such yields results that differ from the prototype. However, the principles are the same, thus the prototype is valuable as a proof of concept.

3.1.3 Buffering and Node parameters

The `RTAP` class handles the input and output signals, along with the timing of the loop, as shown in Figure 3.5. The audio interface fills up its internal buffers and sends it to the operative system of the computer. When the timing indicates that the buffers are full, the `RTAP` reads the operative system buffers for a signal length of L_{signal}, set to be the sample number of the total signal pulse at system sample rate f_s.

In the `Node` class, the new set of buffers are held. These are of length $2 \cdot L_{signal}$, such that a complete pulse signal will always be present in a one of the loop process iterations. Once a new buffer is ready for the `RTAP` class, the `Node` moves the present signal in the buffer L_{signal} backward, and inserts the new `RTAP` buffer.

The principle is illustrated in Figure 3.6. From the figure, it is evident that a signal may be detected up to three times. Half the signal may be detected in the first buffer, and then the full signal in the second buffer, and half the signal again in the third buffer. The buffers in the `Node` is processed further by the subclasses.

3.2 System parameter selection

Based on the available hardware for the prototype, selected system parameters is presented in this section.

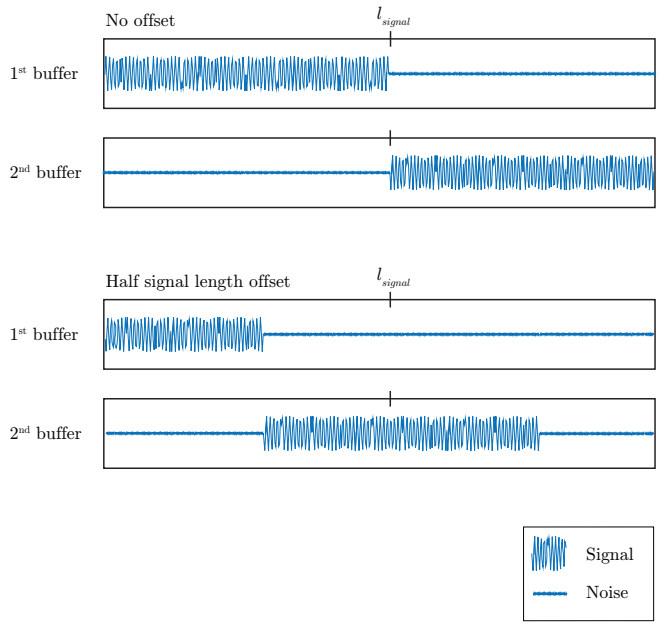


Figure 3.6: Double signal length processing in the `Node` buffers. A complete pulse signal will always be present in one of the consequent buffers.

The first important constraint is the system sample rate f_s. The audio interfaces presented in section 3.1.1.4 supports 96kHz and 192kHz. Assuming the system sample rate to define the minimum time difference detectable by the system, which is true for a correlation where the carrier phase and other advanced timing estimation algorithms are ignored, the detectable time difference is

$$T_{res} = T_s = \frac{1}{f_s} \tag{3.1}$$

The uncertainty in time will be equal to the time resolution as delays of $\pm\frac{\Delta t}{2}$ will be detected as the same time.

Calculating the pulse time difference for $f_s = 96$kHz yields $T_{res} = 10.4\mu$s. For $f_s = 192$kHz, $T_{res} = 5.2\mu$s.

Considering the requirements summarized in section 2.5, the time resolution requirement is $T_{res} = 9.8\mu$s. Thus, the sample rate is selected to be $f_s = 192$kHz such that the pulse time accuracy is within the requirement.

The time uncertainty doubles due to two detector channels such that $\Delta t_\theta = 2T_{res}$. From equation 2.4, the angular error for an on-axis plane wave is given as

$$\Delta\theta_{on-axis} = \sin^{-1}\left(\Delta t_\theta \frac{c_{sound}}{d_{mic}}\right) \approx 0.93° \tag{3.2}$$

This uncertainty is right below the limit of the perceivable change in angle. The error at maximum angle 90° is given as

$$\Delta\theta_{max} = \sin^{-1}\left(1\right) - \sin^{-1}\left(1 - \Delta t_\theta \frac{c_{sound}}{d_{mic}}\right) \approx 10.3°$$

which is slightly above the perceivable angular change for 90° incident sound. Thus the requirement in section 2.5 should be considered fulfilled.

For distance calculations, each of the two nodes' detector introduces the T_{res} error. The error in distance calculations due to the sample rate resolution is thus $\Delta t_x = 2T_{res}$. For a speed of sound $c_{sound} = 343$m/s, equation 2.7 yields a distance uncertainty of

$$\Delta x = \Delta t_x \cdot c_{sound} \approx 3.6\text{mm}$$

It is important to note that these equations are valid only for plane waves and without the extra travel distance associated with the head. Further discussions will be presented in section 4.1.3.

According to section 3.1.1.6 and Figure 3.3, the optimal carrier frequency is 40.8kHz for the selected transmitter. Due to different demodulator implementations which will be described in section 3.5, the carrier frequency should be an integer fraction of the sample rate f_s. Thus, the closest integer fraction is selected, and a carrier rate $f_c = 38.4$kHz was selected. As the frequency is not an ideal match for the transducer, a reduction of the signal level compared to optimum must be expected. However, the output was considered sufficient for measurements and system analysis.

The chip sequence length L_{chip} must be regarded in close relation to the chip rate f_m. The combination of the two parameters affects the noise sensitivity as discussed in section 2.4.6, as well as the transmission time per pulse. Thus, both the refresh rate of positions and the range of the system will be affected.

Considering equation 2.23, the signal-to-noise ratio for the system is given by the detection threshold and processing gain. As the detection threshold is constant, the processing gain is the only adjustable parameter. Assuming the near-far problem discussed in Section 2.4.4.2 is more important than the issue of background noise, PG should be maximized, and the chip rate f_m adjusted later as a compromise between the transmission time and background noise.

The sound intensity difference for two transmitters with distance 4m and 1m to a receiver, respectively, is $20\log_{10}\frac{4}{1} = 12$dB. Thus, if these transmitters send a pulse at the same time, the receiver must be capable of detecting a signal with SNR $= -12$dB to receive the pulse from the node furthest away.

The processing gain can be calculated from Equation 2.23:

$$PG = DT - \text{SNR} \tag{3.3}$$

Inserting $DT = 15$dB and SNR $= -12$dB, $PG = 27$dB. Thus, from the definition of processing gain in equation 2.22, the chip sequence length needed is $L_{chip} = 501$ for the example described above. For larger differences, longer chip sequence lengths are obviously needed.

As described in section 2.4.3, the chip rate defines the bandwidth of the signal. As the transducer has a relatively narrow bandwidth, the chip rate is limited by this constraint. Due to the implementations described later in this chapter, the chip rate must be an integer fraction of the carrier rate f_c. Thus, the following chip rates can be used: $f_m = \{3.2, 4.8, 6.4, 12.8, 19.2\}$kHz.

An experimental approach was used to find the upper limit for the chip rate, restrained by the transducer. Measurements and evaluation revealed that $f_m =$ 6.4kHz should be considered as the upper limit.

Further constraints apply due to implementation-specific details. The decimation processes described in section 3.5 requires an integer fraction of the system sample rate. As well, a hardware-related lower bound for the chip rate is the maximum buffer size $N_{buffer\_max} = 65536$ samples given by the audio interface in combination with the operative system. Forestalling the conclusions made later in this chapter, the final implementation is a dual code system consist of both a preamble- and unique chip sequence. Assuming $L_{chip} = L_{chip,pre} = 511$, with a total chip length $L_{chip,tot} = 1022$, the number of samples for a single transmission is given by

$$N_{samples} = \frac{f_s}{f_m} \cdot L_{chip,tot}$$

Thus, the minimum chip rate is

$$f_{m,min} = \frac{f_s}{N_{buffermax}} \cdot L_{chip,tot} \approx 3\text{kHz}$$

For the double chip length, the minimum chip rate is thus the double.

A link budget for the system with different parameters is presented in Figure 3.7. Here, *low noise level* implies background noise with spectral density -14dB/Hz, 50dB below the *high noise level* with 35dB/Hz as described in section 2.4.6. Two different chip rates, $f_m = 6.4$kHz and $f_m = 12.8$kHz are represented. Several chip sequence lengths are also presented; longer chips implies higher processing gain. Note that directivity is not corrected for; the propagation loss is based on spherical propagation. The SL used in the budget is the measured source level for the directive transducer described in section 3.1.1.6. Thus for the implemented system, the distance will depend on head rotation and decrease for wide angles.

Based on all the considerations above, the chip rate is selected to be $f_m = 6.4$kHz and the chip sequence is selected to be $L_{chip} = 511$. For this parameters, the resulting transmission time is given by

$$t_{transmission} = \frac{N_{samples}}{f_s} = \frac{L_{chip,tot}}{f_m} \approx 160\text{ms}$$

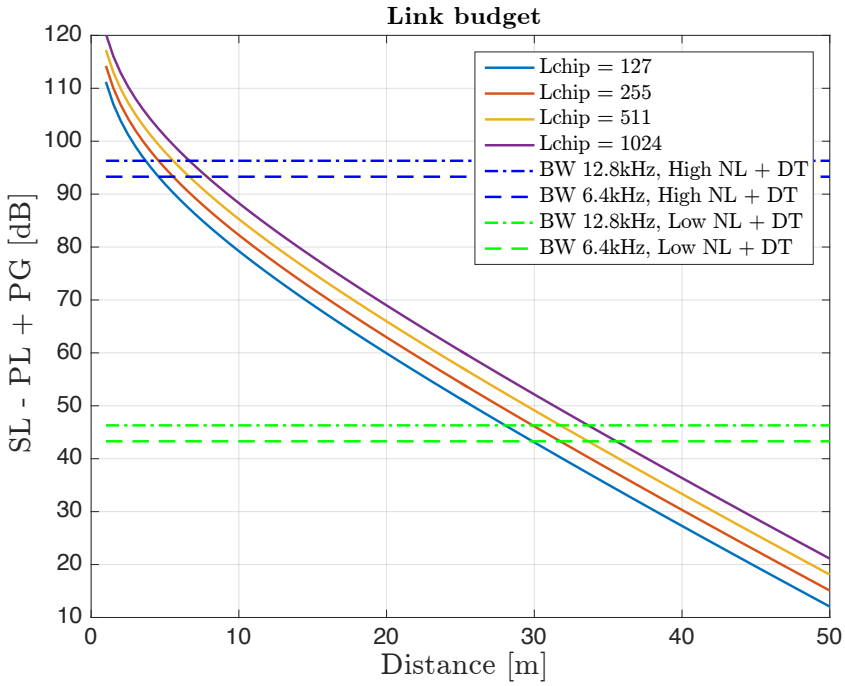


Figure 3.7: Link budget for system implementation, source level 91.5dB SPL, plotted for different code lengths and background noise levels.

As each node requires three codes for identification and commands (discussed in section 3.6.1), the total number of nodes possible in the system is $\frac{L_{chip}-1}{3} = 170$. This number of nodes is unlikely to be present at the same time, and thus, the chip length is not a restriction to the system in this regard.

Considering processing delay and transmission time for three transmissions, the update rate is $3\,(160\text{ms} + t_{processing})$ for the prototype system. In practice, the total time is about 2s. The prototype could have been optimized for a faster update rate, but available time limited the focus on the issue.

The budget in Figure 3.7 suggests that for low-noise conditions, distances of above 30m may be achieved. For high-noise conditions, the system may be restricted to around 7m. Improvements will, of course, be obtained with higher source levels, and as suggested in section 3.1.1.6, the transmitter may be capable of higher levels. Equation 2.27 and 2.28 states the requirements for the transducer with respect to the chip sequence length and chip rate.

A summary of the selected parameters is presented in Table 3.1.

Table 3.1: Summary of system parameters and specifications

Parameter	**Symbol**	**Value**
System sample rate	f_s	192kHz
Carrier rate	f_c	38.4kHz
Chip rate	f_m	6.4kHz
Chip length	L_{chip}	511
Transmission time	$t_{transmission}$	160ms
Angle accuracy 0°	$\Delta\theta_{on-axis}$	±0.47°
Angle accuracy 90°	$\Delta\theta_{max}$	±5.2°
Distance accuracy	Δx	±1.8mm
Range low-noise	x_{max}	34m
Range high-noise	x_{max}	7m

3.3 Spreading code

The pseudo-noise spreading codes selected for DSSS operation was of Gold type and was discussed in section 2.4.3. For the implementation the codes was generated using the *Communications Systems Toolbox* in MATLAB, with a specific function called `comm.GoldSequence`. The function returns a binary data stream of zeros and ones.

The reader is reminded that no data is to be transferred over the system other than the chip sequences. These chip sequences represent the identity of each node, and also what kind of command that is to be sent between the nodes. For the implementation in section 3.5.3, a common preamble sequence is added to the signal before the unique sequence. The requirement for identity and messages is

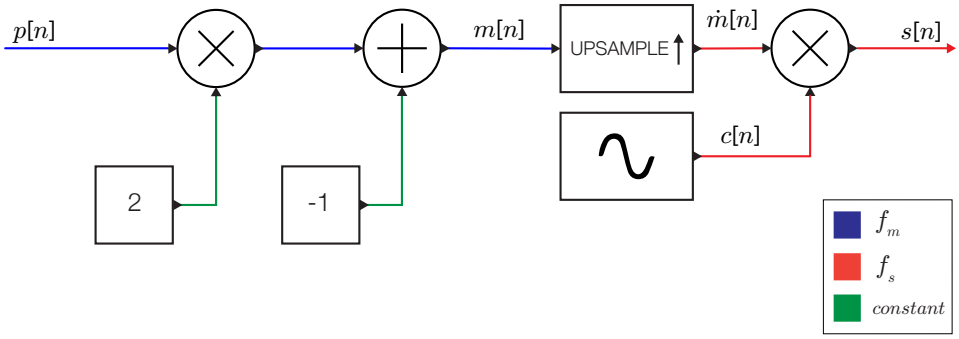


Figure 3.8: SFG for the modulation process

described in Section 2.4.5, while the implemented coding scheme is outlined in section 3.6.1

3.4 Modulation

The modulation process for a DSSS signal using BPSK without any data content is discussed in section 2.4.4.1. In the implementation, three stages of processing are required to accomplish the modulation. A signal flow graph (SFG) for the modulator is shown in Figure 3.8.

First, the spreading code sequence $p[n]$ is multiplied by two and subtracted by one to achieve a phase signal $m[n]$ with values of either 1 or -1. The phase signal is now in baseband at the chip rate f_m.

To allow for modulation of the phase signal onto a carrier, up-sampling to full sample rate is performed by a simple algorithm. The process inserts samples between each of the samples in $m[n]$ and assigns all the new samples the same value as the sample before the new block of inserted samples. The up-sampled phase signal is denoted $\dot{m}[n]$.

The phase signal is thereafter multiplied by a carrier signal $c[n] = cos(2\pi f_c \frac{n}{T_s})$ to up-mix the spreading code to the carrier frequency, where $T_s = \frac{1}{f_s}$.

3.5 Demodulation and detection

The background on demodulation of a burst DS-CDMA signal modulated with BPSK is discussed in section 2.4.4.2. The input is a modulated passband signal, and it is most efficiently processed in baseband. To move the signal down to baseband, a multiplication with a local carrier is performed. A low-pass filter is

applied to the signal after the multiplication to avoid spectral images. After the down-mixing process, the bandwidth of the signal is reduced to the chip rate, and a decimation process may be applied to the signal, thus reducing the computational load on the correlators. The main challenges in the demodulation process are the synchronization of the carrier phase and the chip decision points, and these issues will be discussed thoroughly.

Two different approaches to the demodulation process will be presented. The first approach, named the *Carrier-lock demodulator*, is a conceptually simple approach used to test the overall system without developing and investigating more advanced solutions. However, the solution proved insufficient, and a second approach, *Complex baseband demodulation*, was used in the final prototype implementation.

For clarity throughout the following text, the output of the detectors is a time offset constant with unit samples, at system sample rate f_s. The constant is denoted Δ_M and represents the *lag* of the incoming system. Lag $\Delta_M = 0$ implies that the received signal contains a chip sequence beginning at the first sample of the buffer.

3.5.1 Carrier-lock demodulator

This implementation seeks to solve the carrier phase synchronization by tracking the phase of the incoming signal in each buffer. After phase tracking, a cosine waveform is generated for demodulation. Down-mixing is performed, and the chip decision points are selected based on the zero-crossing detection.

The implementation is divided into several blocks, and the overview signal flow graph is shown in Figure 3.9. First, the *start index locator* block processes the input signal $r[n]$ to estimate the first index i of present signal on the input. Second, the index is passed on to the *Costas loop* block where phase tracking is performed on a limited length of $r[n]$. The last block is the *demodulation* block, performing down-mix and decimation of input signal $r[n]$ along with decision point detection and decimation. The output of the carrier-lock demodulator is a chip rate sequence $m_r[n]$ for further processing in a correlation detector, along with an offset lag constant Δ_{dec} to correct the output lag of the correlator detector.

3.5.1.1 Start index locator

The start index locator is an algorithm that finds evidence of the carrier frequency in the spectrum of $r[n]$, and thereby finding a window in the buffer where it is plausible that a signal is present. The method employs a number of FFT's taken on small sections of the buffer, where the ratio of the carrier frequency power to the mean signal power is used as the indicator. If the carrier to mean ratio is above a certain threshold for a given index i_0, the index is selected, and a signal is assumed present from the index of the window location. One can easily see that this method requires a positive signal-to-noise ratio to function as intended. The signal flow graph of the process is presented in Figure 3.10.

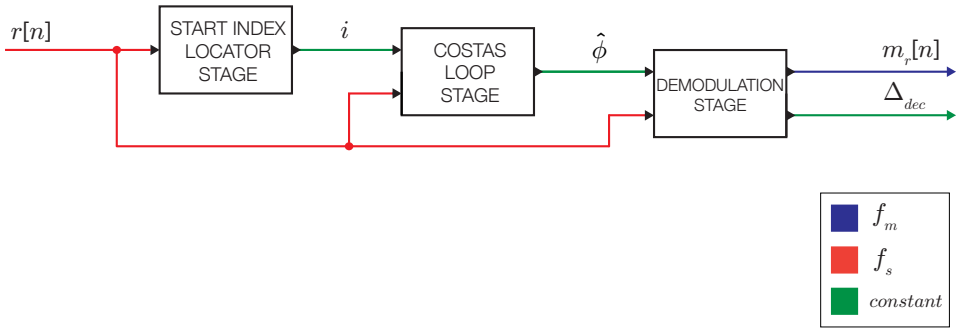


Figure 3.9: Carrier-lock demodulator overview

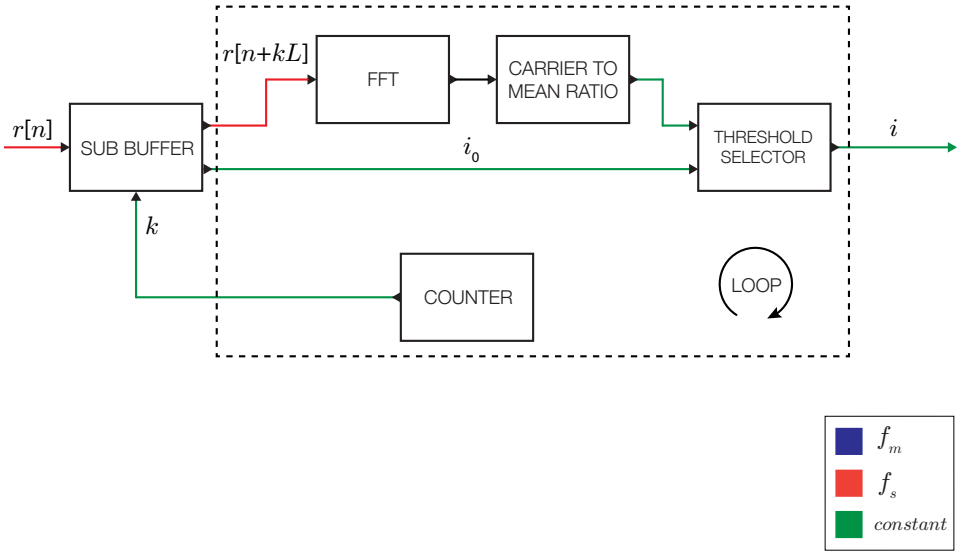


Figure 3.10: SFG for the start index locator stage in the Carrier-lock demodulator

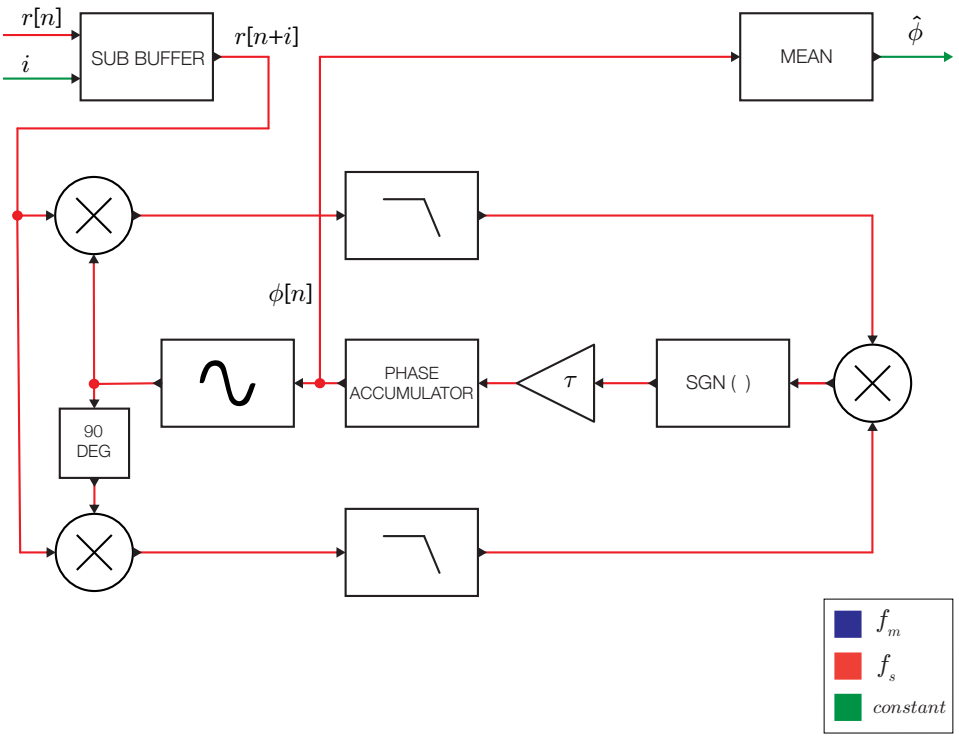


Figure 3.11: SFG for the costas loop stage in the Carrier-lock demodulator

3.5.1.2 Costas loop

The Costas loop block is an application-specific implementation of a phase locked loop for communication systems. The principle is mentioned in section 2.4.4.2, and a modified example code from [37] is used in the implementation. A sub-buffer feeds the signal $r[n+i]$ with i obtained from the start index locator for the known signal length L_{signal} into the algorithm. Tuning of the feedback loop sensitivity constant τ is necessary to achieve a suitable phase convergence[3]. The last part of the phase output $\phi[n]$ is averaged to obtain a stable phase offset and reduce noise interference, and the estimated phase $\hat{\phi}$ is the result of process. The signal flow graph of the Costas loop implementation is presented in Figure 3.11.

3.5.1.3 Demodulation

The main block in the Carrier-lock demodulator is the demodulation block. The signal flow graph is illustrated in Figure 3.12. The phase $\hat{\phi}$ obtained from the Costas loop is utilized to generate a local carrier for multiplication with the incoming

[3]The constant was experimentally set to $\tau = 10^{-3}$s

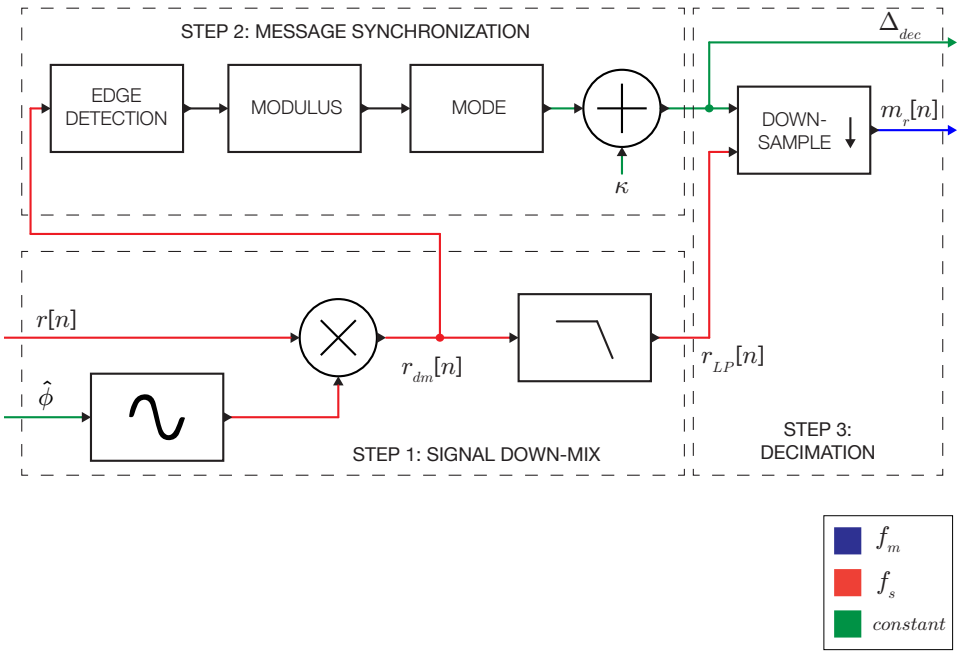


Figure 3.12: SFG for the demodulation stage in the Carrier-lock demodulator

signal. After the multiplication is performed, a low-pass filter is applied[4], and the resulting down-mixed signal is now at chip rate f_m and may be decimated. To obtain the highest possible signal-to-noise ratio, the decision points needs to be selected carefully.

To find the decision point, an edge-detection algorithm is used. This algorithm samples the zero crossings in down-mixed signal $r_{dm,LP}(n)$, takes the modulus of these samples with respect to the chip rate and calculates the mode[5] of the positions to improve the accuracy.

After the edge detection is accomplished, a constant $\kappa = \frac{1}{2} \cdot \frac{f_s}{f_m}$ is added, thus targeting the peaks of the low-passed signal. The peaks are down-sampled to the message rate at the decision points, yielding the received message (the PN-code) $m_r(n)$ along with an offset scalar Δ_{dec} to allow for lag compensation due to message synchronization in the detector. The implementation may benefit from replacing the decimation with an integrate-and-dump technique, but this has not been implemented and tested.

[4]Designed by the `firceqrip` MATLAB function, $f_{passband} = f_m/6$, ripple pass-band $rp = 0.001$, ripple stop-band $rst = 1E-6$, order $N = 100$. The filter is steep to provide an integration-like signal before decimation.

[5]the number occurring most frequently

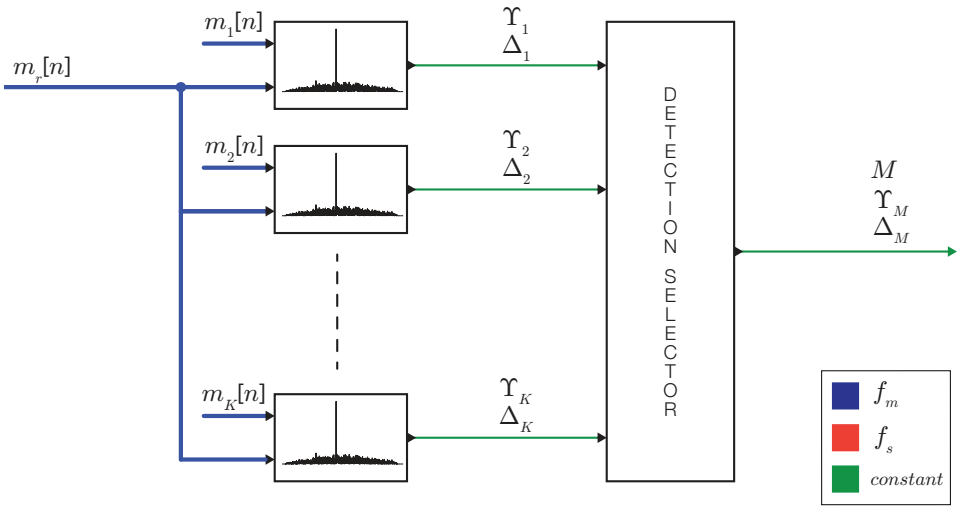


Figure 3.13: SFG for the carrier-lock correlator bank

3.5.2 Detector for the Carrier-lock demodulator

The detector contains the processing elements for code detection and calculation of code lags. The down-mixed signal $m_r[n]$ is correlated with each of the code sequences $m_k[n]$ to produce a correlation function $R_{m_r,m_k}[n]$ for all codes $k = 0$ to $k = K$. A search for peak coefficients is performed and the first index n that fulfils the requirement

$$\frac{R_{m_r,m_k}[n]}{R_{m_r,m_k}[n-1]} > \Upsilon_{threshold}$$

is taken as a detected code. The threshold is set experimentally to allow for maximum detection capability without false detection. All level indicators Υ_k is passed on to a selector algorithm which selects the largest level as a detected code, given a positive level indicator from at least one correlator. The selector outputs the level indicator Υ_k, the lag Δ_k and the selected code M for further processing in the system.

3.5.3 Complex baseband demodulation

The detection performance and distance results of the Carrier-locked demodulator proved to be sufficient for noise-free environments, but measurements revealed that angle calculations gave significant inaccuracy and spread. Occasionally, the signal was detected by only one channel, and in noisy conditions, the detection would fail. Simulations and investigations indicated that failure to synchronize carrier and chips were the issue.

To improve the carrier synchronization issue, a complex baseband demodulation approach was investigated and implemented. The principle is described in section 2.4.4.2. In this approach the carrier synchronization is omitted; the cosine carrier is replaced by a complex carrier, such that a complex baseband demodulation is achieved. This is equivalent to two-branch processing where one branch is the input signal multiplied with a cosine, and the other branch is the input signal multiplied with a cosine 90° offset to the other. This is commonly denoted IQ-processing.

The other main issue is the chip synchronization. In the Carrier-locked approach, edge detection was exploited. However, when subject to noise this method is not reliable. This will also be the case in the complex baseband approach. Several solutions to the problem have been investigated, but considering limited time and resources two methods crystallized. These will be presented below. Both methods may, in theory, work with the same coding scheme as described earlier, but in practice, the computational abilities of the prototype real-time system resulted in overloading.

To mitigate the computational overload issue, an extended signal scheme is used. By doubling the transmission time, a synchronization code sequence is added before the unique code. The synchronization sequence is for further reference named the *preamble*. The preamble is from the same Gold code family as the unique code set, and with the same chip length $L_{chip,pre} = L_{chip}$ to maintain a common chip generator and the same processing gain. The preamble is common to *all* nodes, and is used for precise synchronization and calculation of TOA for the incoming signals, while the unique code is used for node identification and logic. In a multi-user perspective, several positive detections may be output from the preamble correlation as the preamble is shared among all users. The TOA for the unique chip sequence must be in coherence with the TOA from the preamble correlation (within a precision interval) to ensure that the preamble TOA belongs to the detected unique code. The two new strategies for synchronization are:

Synchronization correlator bank The preamble is synchronized by a dedicated correlator bank using different decimation offsets. Requires N correlators for synchronization, where $N = \frac{f_c}{f_m}$ represents the number of possible integer decimation offsets in a carrier. The method is presented in section 3.5.3.1.

Full-rate synchronization The code is synchronized before decimation to obtain a sample-rate resolution for the lags. This method is presented in section 3.5.3.2.

Cross-correlations are computed by FFT for lengthy signals in MATLAB. In an

embedded system, the same method may be applied, but this will require a buffer design. A general estimation of the number of operations required to compute the correlation is

$$O(N \log(N)) \tag{3.4}$$

for a N-length signal.

Considering the possible strategies for implementation of the correlator code, calculations of the computational load may be estimated. For the synchronized correlator bank without preamble, the number of operations will be

$$O = N_{nodes} \cdot \frac{f_c}{f_m} \cdot L_{chip} \cdot \log\left(L_{chip}\right) \tag{3.5}$$

For the same system with preamble, the number of operations is

$$O = N_{nodes} \cdot L_{chip} \cdot \log\left(L_{chip}\right) + \frac{f_c}{f_m} \cdot L_{chip,pre} \cdot \log\left(L_{chip,pre}\right) \tag{3.6}$$

The full-rate synchronized system without preamble will obtain a number of operations as

$$O = N_{nodes} \cdot \frac{f_s}{f_m} \cdot L_{chip} \cdot \log\left(L_{chip}\right) \tag{3.7}$$

and with preamble

$$O = N_{nodes} \cdot L_{chip} \cdot \log\left(L_{chip}\right) + \frac{f_s}{f_m} \cdot L_{chip,pre} \cdot \log\left(L_{chip,pre}\right) \tag{3.8}$$

To study the impact on the real-time system, consider an implementation with the following key parameters:

Sample rate	$f_s = 192\text{kHz}$
Carrier rate	$f_c = 38.4\text{kHz}$
Chip rate	$f_m = 6.4\text{kHz}$
Chip length preamble	$L_{chip,pre} = 511$
Chip length unique	$L_{chip} = 511$
Total chip length	$L_{chip,tot} = L_{chip,pre} + L_{chip} = 1022$
Number of nodes	$N_{nodes} = 12$

Using the lowest computation as a reference, the ratios of the different implementations in our example yields the following result:

System	**Without preamble**	**With preamble**
Sync. correlator bank	4.0	1.0
Full-rate sync.	30.1	3.2

It is evident that the preamble drastically reduces the number of operations required to demodulate the signal. The synchronized correlator with preamble stands out as the algorithm with the least number of operations. The actual gain of using a preamble code will, of course, increase with the number of nodes.

The preamble scheme is chosen as the strategy in the implementation of the Complex baseband. It will later be shown that the full-rate implementation is superior to the message rate correlator bank, which will only be briefly discussed. The trade-off is a somewhat higher computational load, but it is still within the limitations of the real-time system.

3.5.3.1 Chip-rate preamble IQ correlator bank

The chip-rate preamble correlator bank strategy is the most computational efficient strategy. The main idea is that a bank of correlators all processes the same input signal and chip sequence, but with different decimation points. Choosing the best correlation corresponds to selecting the correct decision point. Parallels may be drawn to the RAKE receivers [20, sec. 13.2.4] designed to handle multipath signals in telecommunication, but with another purpose. The signal flow graph of the implementation is presented in Figure 3.14.

The down-mix process comprise the same basic architecture as the Carrier-locked implementation, but the real carrier signal $c[n]$ is substituted with a complex carrier $\boldsymbol{c}[n]$, resulting in a complex baseband signal $\boldsymbol{r}_{LP}[n]$ after a complex low-pass filter process. The preamble correlators operate on this complex signal, thus resulting is a complex correlation coefficient sequence. To obtain the peak value, the absolute value of the correlation output is exploited. The output from each of the N correlators is the maximum correlation level $\Upsilon_{pre,n}$ and the lag $\Delta_{pre,n}$.

During implementation, simulations revealed that when exposed to noise, the highest correlator output does not necessarily correspond to the correct decision point when considering timing. In a communication system this selection would indeed be satisfactory, but in a lag-critical application, the brute selection of synchronization causes unstable lag results and insufficient precision. Thus, a simple $\max(\Upsilon_{pre})$ algorithm for the decimation selector is not sufficient. Instead, the selector exploits the difference in output lag $\Delta_{pre,n}$ to seek transitions, and chooses the decimation point to be in the estimated center of the chip. The selector then outputs the decimated complex baseband signal $\boldsymbol{m}_r[n]$ based on the decimation point selection for further processing in the unique chip sequence correlator bank along with a corresponding offset lag constant Δ_{pre}.

The unique chip sequence correlator bank identifies the purpose of the signal and calculates the precise lag of the incoming signal. The maximum peak is found from the absolute value of the correlation function $R_{\boldsymbol{m}_r,m_k}[n]$. If the peak is above a threshold level Υ_{thres} a code is considered detected. The threshold level is tuned experimentally to neglect false code detections, but still retain maximum performance at low signal to noise ratios.

The coarse lag for a given code k is found as the lag output Δ_k plus the decision correction Δ_{pre}. Precision is still lacking as the decision point correction does not compensate for the phase offset of the incoming signal carrier. The fine lag is calculated by unwrapping and extracting phase offset from $R_{\boldsymbol{m}_r,m_k}[n_{peak}]$ and the neighbouring sample as described in [26]. The total lag output of the correlator

bank Δ_M is thus the sum of the coarse lag and the fine lag.

The detection selector takes the detected code M as the correlator with the highest detection level output Υ_k. To deal with the issue of multiple detections of signals partly outside the buffer (due to the overlap as described in section 3.1.3), the detection is neglected if the lag indicates that the signal or preamble is outside buffer.

3.5.3.2 Full-rate preamble IQ correlator

The full-rate preamble correlator takes a simpler approach to the decision point synchronization issue. The signal flow graph for the implementation is presented in Figure 3.15. Instead of decimating the preamble signal, an up-sampled preamble code $\dot{m}_{pre}[n]$ is correlated directly with the down-mixed complex baseband signal $\boldsymbol{r}_{LP}[n]$. The main advantage of this design is the accurate lag output Δ_{pre} from the correlator, without any phase π-ambiguity or decimation point selection error. Thus, an accuracy $T_{res} = \frac{1}{f_s}$ is achieved as discussed in Section 3.2. The lag output is further used to select the correct decision point for decimation and processing in the unique code correlator bank by a simple modulus operation in the detection selector.

In the full-rate preamble correlator, the preamble lag Δ_{pre} is used for the detection of lag, and the unique code correlator bank is used merely for identification of node and purpose of the signal. The correlator principle is the same as described for the chip-rate preamble detector. The detection selector uses the preamble lag as the code lag for the code with the highest detection level Υ_k. The detection selector also prevents multiple detections due to lag outside the buffers in the same manner as the chip-rate preamble correlator.

3.5.4 Detector performance evaluation

To quantify the performance of the demodulator implementations, simulations were used. The signal to noise ratio is varied using worst-case noise; irrelevant modulated chip sequences are summed and modulated to create a noise spectrum equivalent to the signal spectrum. A modulated chip sequence of interest is generated, and the noise signal is added to the modulated chip sequence before demodulation.

Along with variation of noise level controlled by the number and scaling of irrelevant codes added to the signal, the time offset of the input chip sequence is varied in the performance simulator. This prevents the synchronization algorithms to have similar working conditions for all iterations and introduces a natural variability of the input.

The signal is considered detected if the detector has successfully detected the sequence of interest along with a lag result which is within a certain interval relative to the input lag. Two cases were simulated, a ± 20 lag interval and a ± 5 lag interval. The number of iterations for each SNR value is set to 500, and the output is therefore capable of measuring errors rates to a probability of $\frac{1}{500} = 2 \cdot 10^{-3}$. The

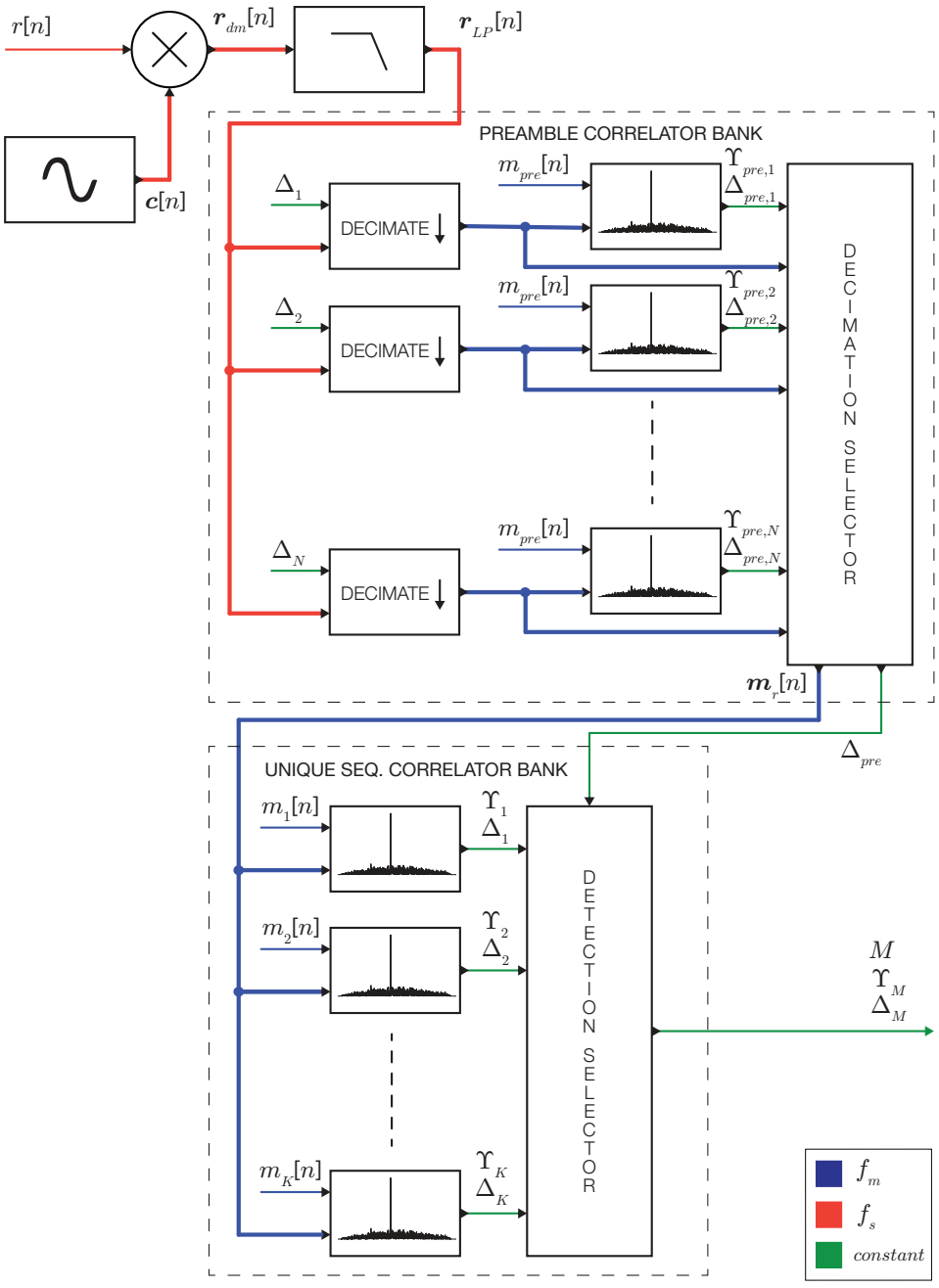


Figure 3.14: SFG for the chip-rate preamble IQ demodulator. Complex signals are denoted with bold symbols lines.

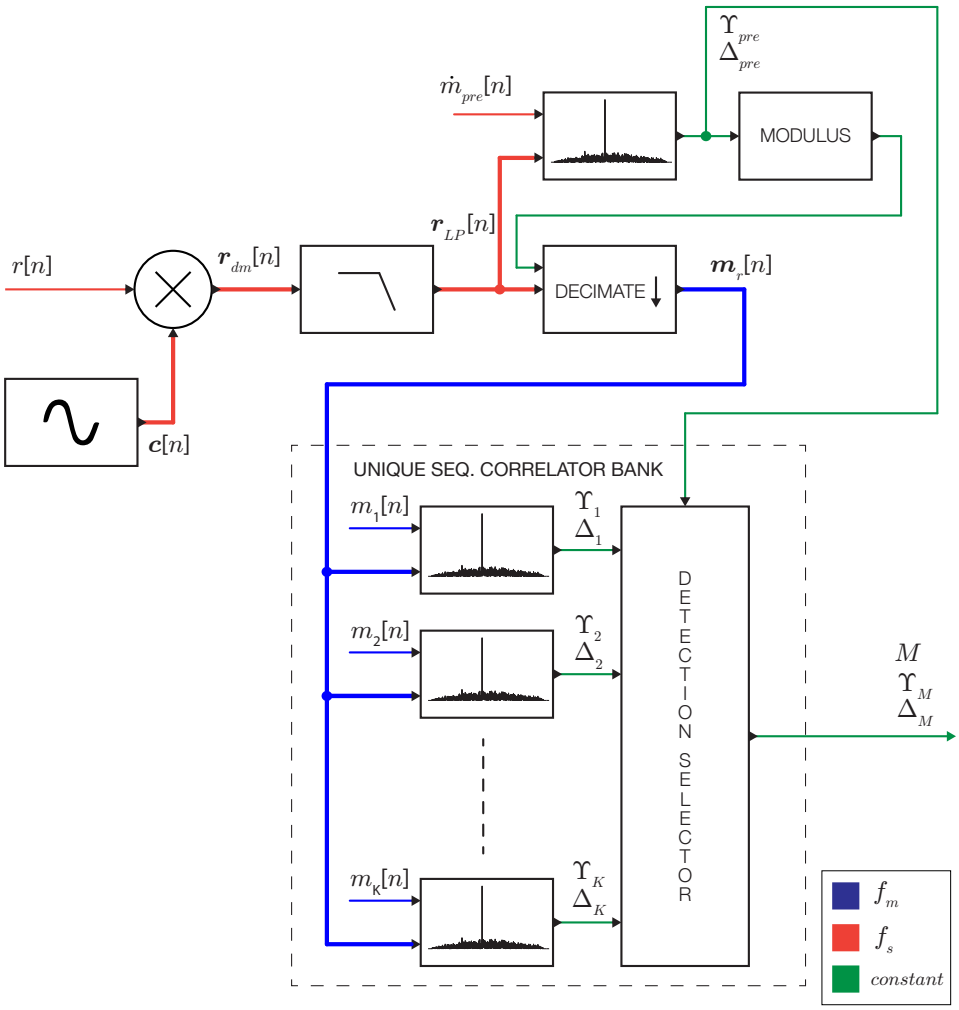


Figure 3.15: SFG for the Full-rate preamble IQ demodulator. Complex signals are denoted with bold symbols lines.

simulations are only performed for negative SNR values. The results are presented in Figures 3.16a and 3.16b.

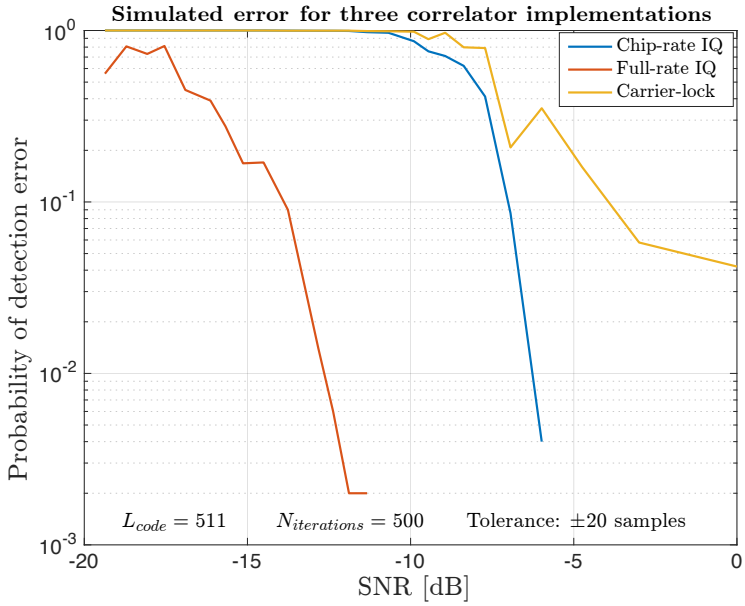
The results show that both of the IQ-based preamble detector implementations clearly outperforms the Carrier-lock detector for both lag thresholds. As well, it is evident that both the Carrier-lock implementation and the Chip-rate preamble IQ implementation has limitations to the precision of lag detection. The full-rate preamble IQ implementation does, however, provide stable results for both lag detection intervals, proving that the chip synchronization technique using a single full-rate correlator is superior to the synchronization technique in the chip-rate preamble IQ demodulator implementation.

If the lag precision is treated as a secondary concern and focus is on the ability to detect the presence of a signal, the graph in Figure 3.16a shows that the chip-rate preamble IQ implementation may handle SNR of about -6dB, while the full-rate preamble IQ implementation may handle SNR of around -12dB. Considering lag precision, only the full-rate IQ implementation shows significant performance in the negative SNR range. The error rate is zero all the way down to -12dB SNR.

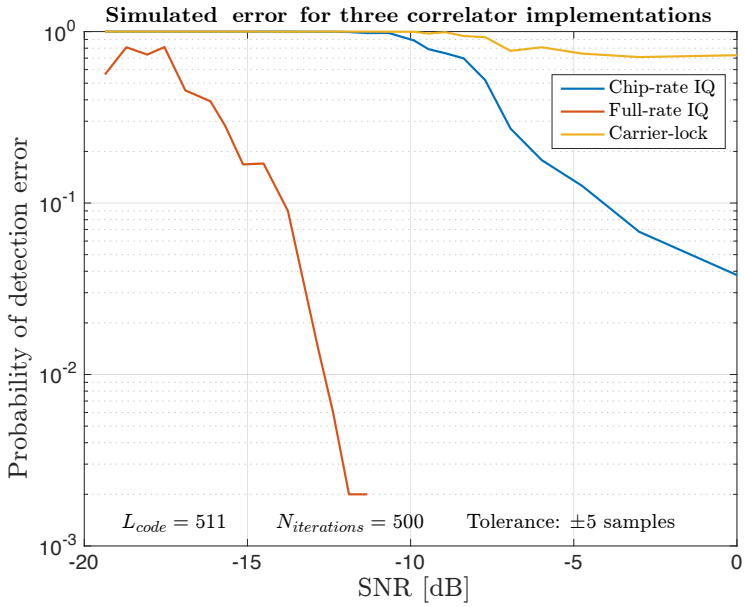
These number corresponds well with the SNR values presented in 3.2. For $L_{chip} = 501$, close to the actual chip sequence length, the expected SNR is -12dB for $DT = 15$dB.

An important note is the capability to handle CDMA properly. Neither of the classes has been implemented to exploit the full capacity of a CDMA system. The carrier-lock demodulator is only capable of synchronizing to a single carrier inside each buffer. The most powerful signal will be synchronized, but multiple signals will most likely jam the synchronization if the signal power difference is small.

The baseband demodulators are capable of handling multiple overlapping signals to a much larger degree. The chip-rate preamble correlator needs a more complex algorithm to do this, as different lags must be handled separately for each possible incoming code. For the full-rate preamble correlator, the implementation is rather easy. All lags above the detection threshold should be matched separately with the lag outputs from the unique correlator bank, and sorted as such. This implementation is not considered very advanced, and lack of implementation is only due to the time constraints of this work. Based on the discussion above, the full-rate preamble correlator seems to be the best solution with respect to CDMA as well.



(a) Detection error rate for detector implementations with lag interval ±20



(b) Detection error rate for detector implementations with lag interval ±5

Figure 3.16: Detection error rate simulations for correlator implementations.

3.6 Signal transmission and positioning logic

The logic architecture is described in this section. As discussed in section 2.4.5, the unique sequence scheme is is based on a broadcast-response-retransmission scheme, requiring CDMA operation. Thus, collisions problems are assumed to not be an issue outside the near-far problem. Each code is implemented as PN-sequence from the Gold code family, discussed in sections 3.3 and 2.4.5.

3.6.1 Signal command scheme

Each node is assigned an identification number. Three types of commands belong to each node. The Request command Q_k is used to broadcast that a node wants a response, where k is the ID of broadcaster. Other nodes respond to this request by the response command S_k where k is the ID of the responder. The last command is the retransmission command, R_k, issued by the initial broadcasting node after reception of a S_k. Here, k is the ID of the responding node who sent S_k. This way, other nodes may ignore the retransmission command. The command scheme is summarized in table 3.2, where the actual sequence number is shown. As each node has three associated sequences, a multiplicand of three is used.

Table 3.2: Chip sequence scheme

Code #	**Function**	**Sent by**
$3 \cdot ID_{request\ node}$	Request	Requesting node
$3 \cdot ID_{response\ node} + 1$	Response	Responding node
$3 \cdot ID_{response\ node} + 2$	Retransmission	Requesting node
$L_{chip} - 1$	Preamble	

3.6.2 Command reception logic

To handle the output of the detector processes, two logic processes is implemented as classes. These classes do not handle signals, but rather logic input variables. The first class, the `Timekeeper`, handles takes the input from the `Detector`. The `Timekeeper` is not complex process. The main purpose is to sort the incoming detection matrix into a variable-length output with the lags Δ_m for each sequence number m added to the total absolute time (in samples). The process thus handles the system time, counted in samples[6].

The second class is the `Communicator`. This process handles the communication logic as well as calculations of distance and angle. The sorted output from the `Timekeeper` is processed for each code detected. Registers are implemented to keep track of expected input codes, last update times for each node and pending

[6]The first iterations of the prototype implementation utilized the Carrier-lock demodulator, and the output offset from the demodulator was handled and corrected by the `Timekeeper` class.

output codes. Pending output codes are passed to the `Modulator` when the schedule is reached. The process also handles periodic requests and data storage for measurements.

An incoming command loop is the first step. For each update from the detector, a range of logic decisions are made. The process is best described graphically, illustrated in Figure 3.17. The registers are first checked for whether a command pulse should be expected, or if the node recently has updated. Thus, irrelevant or redundant commands can be prevented, reducing the traffic on the ultrasonic channel. If a command is expected, different routines apply to various command types. Absolute arrival time is stored for the corresponding node and the queues for processing later in the algorithm is set.

The next step is a node update loop. The algorithm calculates angles and distance based on the arrival times. The process is illustrated in Figure 3.18. The algorithm checks if the lag difference for each channel corresponds to a physical angle value. If not, the first incoming lag is selected, and the corresponding channel is selected as the input. The angle is set to $\pm 90^\circ$ depending on which channel that obtains the first incoming command pulse.

After the two loops described above are completed, the last part of the process is the output queue loop. The algorithm loops through the output queue for each node and checks whether a command should be sent to modulation. If a pending absolute lag is outside the current buffer, no action is performed. Otherwise, the buffer-relative start sample for the output pulse is calculated, and the `Modulator` is signalled.

3.7 Positioning with IMU

The final process is the `Positioner` class. The purpose of this class is to track the IMU signal and he positions of other nodes relative to itself. In the implementation, focus is on presenting the positions in a plot with a static room reference.The plot is calibrated upon initialization, and head rotation angle from the IMU data is presented along with the positions of other nodes. The plot is updated for each buffer process in the `RTAP` class. An illustration of the plot is presented in Figure 3.19. In addition, positions are stored and printed to the command line if needed. For subjective evaluation of the system, the data may easily be output to an HRTF processing module, either internally in MATLAB or via an interface.

3.8 Measurement-specific implementation

A number of features are implemented solely for the purpose of verification and testing. These implementations are listed here to provide an understanding of how measurements and results are generated.

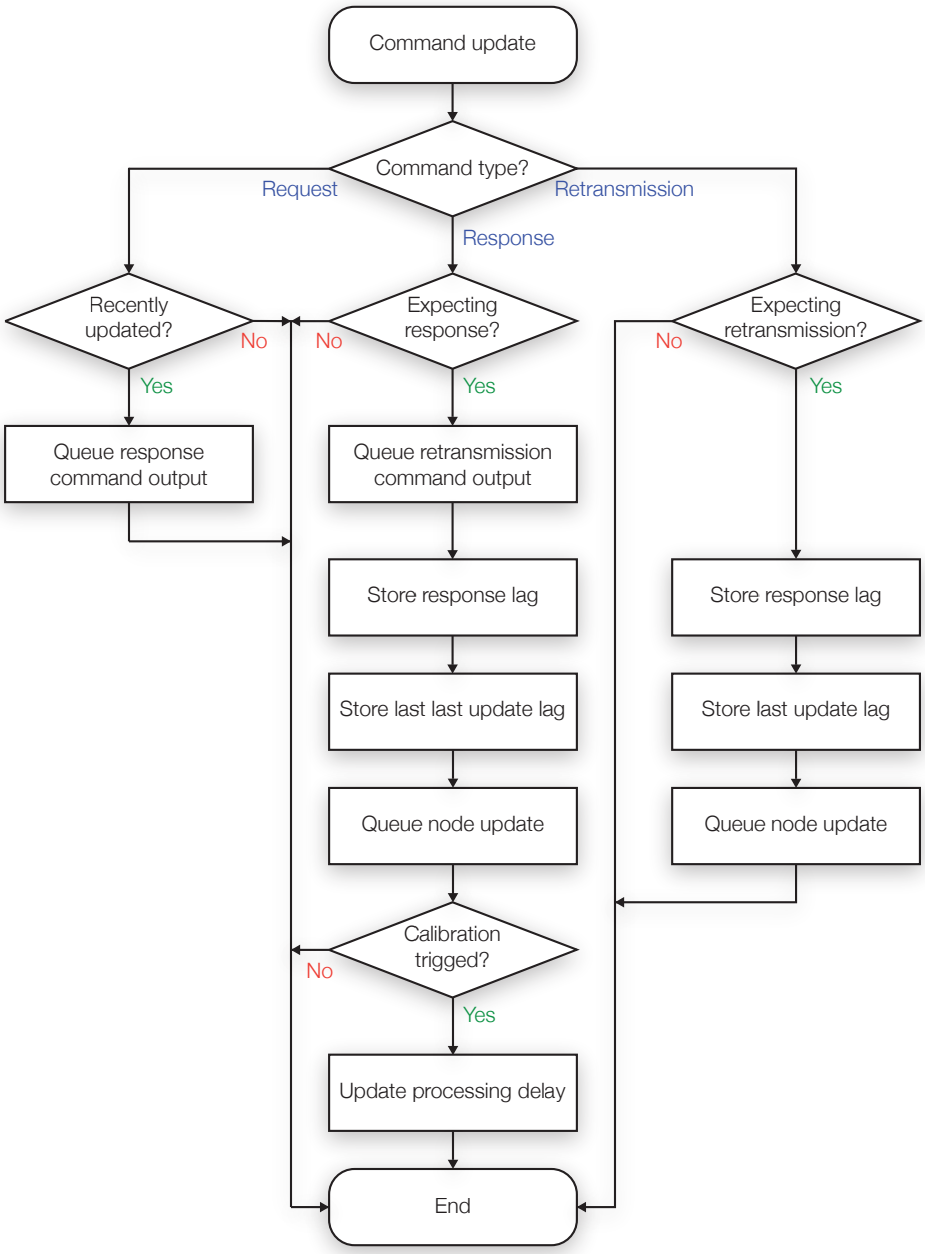


Figure 3.17: Flow chart for the received command logic in the Communicator process

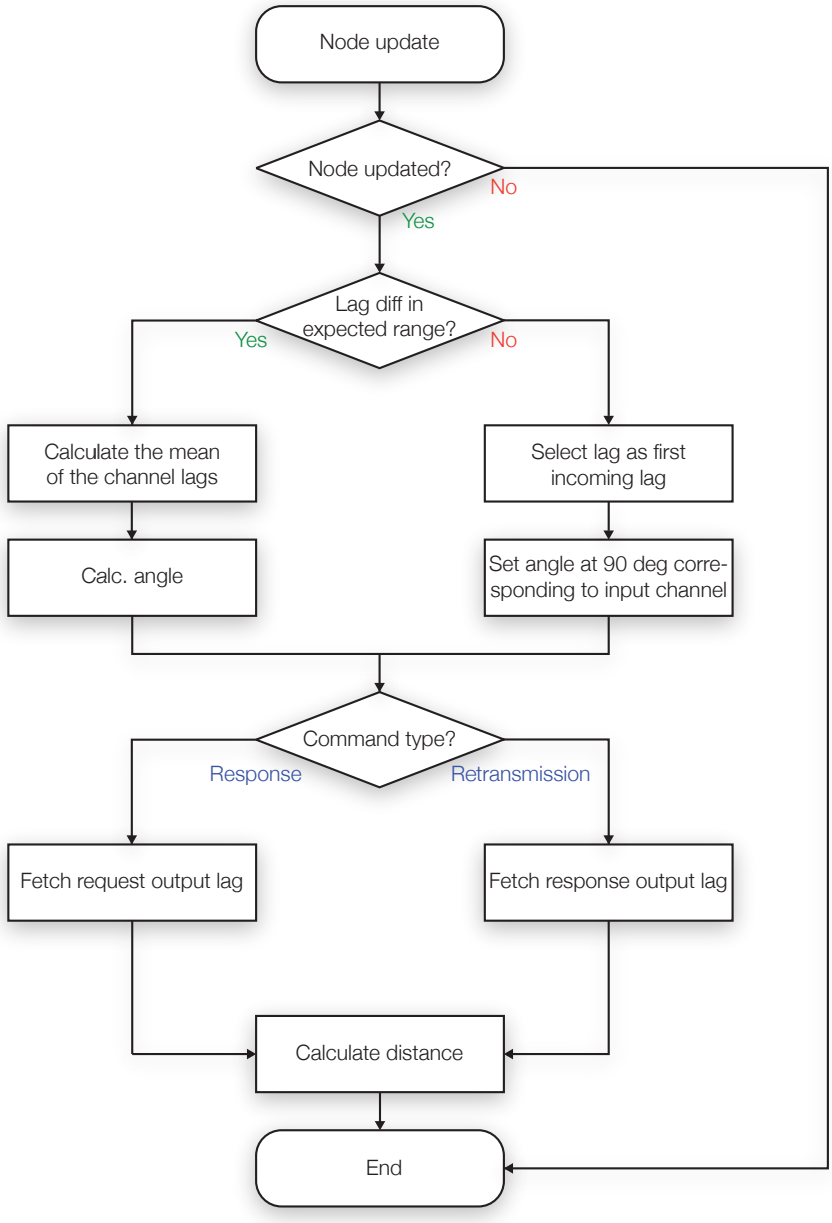


Figure 3.18: Flow chart for the node update logic in the Communicator process

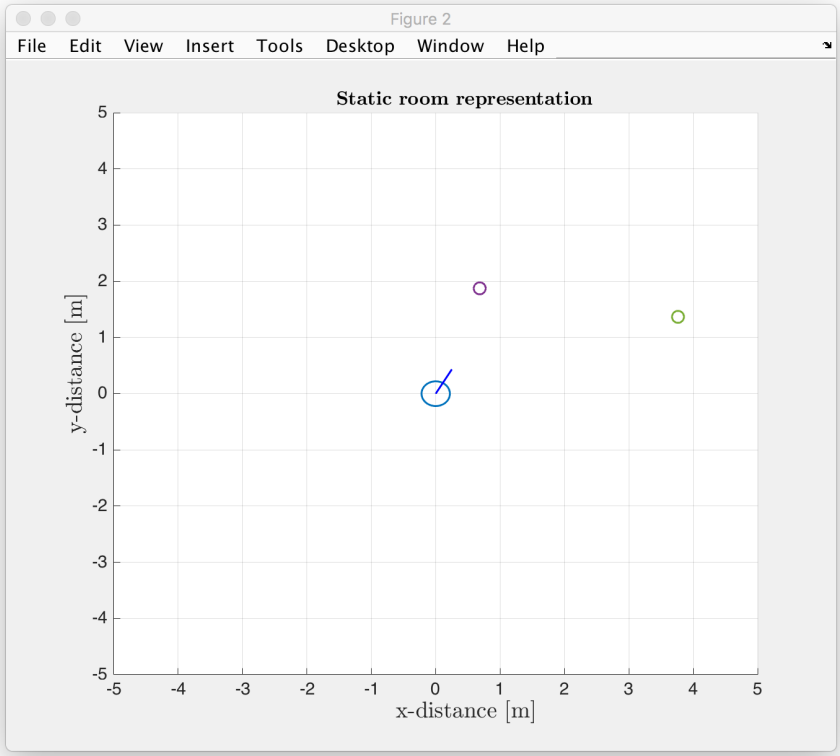


Figure 3.19: Screendump of the position plot presented by the Positioner. The IMU data has reported a 30° angle, shown by the blue line in the centre of the circle representing the node under test. Node 2 (purple) is at 2 meter and $-10°$, Node 3 (green) is at 4 meter and 40°

Turntable control
The `RTAP` class includes a `Turntable` class to control the turntable via a serial port. The `RTAP` GUI supports change of angle, and the class has configurable parameters for speed and acceleration. As well, the class supports continuous rotation mode, allowing the turntable to rotate continuously between two boundary angles.

Periodic transmission mode
The `Communicator` contains an implementation that allows for periodic transmission of request signal. This feature is the preliminary implementation of the normal transmission behaviour of the system. The implementation contains a random time function that introduces a variable offset of the output signal in the buffer, such that two systems will not be perfectly synchronized with respect to the position a received code will obtain in a buffer. The periodic transmission may be switched on and off in the `RTAP` GUI. The `Communicator` also stores all received and calculated values appearing during the periodic transmission to generate result data for analysis. The data stored is

- Node ID
- Calculated distance
- Calculated angle
- Lag difference between channels

Attenuation
To enable signal-to-noise ratio measurements, an attenuation option is available through the `RTAP` class. The attenuation is set in decibels in the GUI and affects the output level via a multiplication performed on the output buffer.

Level printout
The `Detector` has a level read-out option inside the loop, such that the RMS-level of the incoming signal is printed to the MATLAB command window during operation. This feature allows the input level from different sources to be calibrated to the same level for SNR measurements. The function is activated manually in the `Detector` class.

Level calibration
A level calibration function is implemented to provide a steady signal output for calibration of signal-to-noise ratio measurements. When activated, the modulator provides a continuous output of the preamble code, such that the signal level may be detected by another node via the *Level printout* function. The combination of these two functions also allows for gain calibration for the audio interface and microphones.

Chapter 4

Measurement and Verification

This chapter discusses the measurement setups and present results for several tests of the system. Important parameters and implications will be discussed, and possible solutions will be proposed at the end of the chapter.

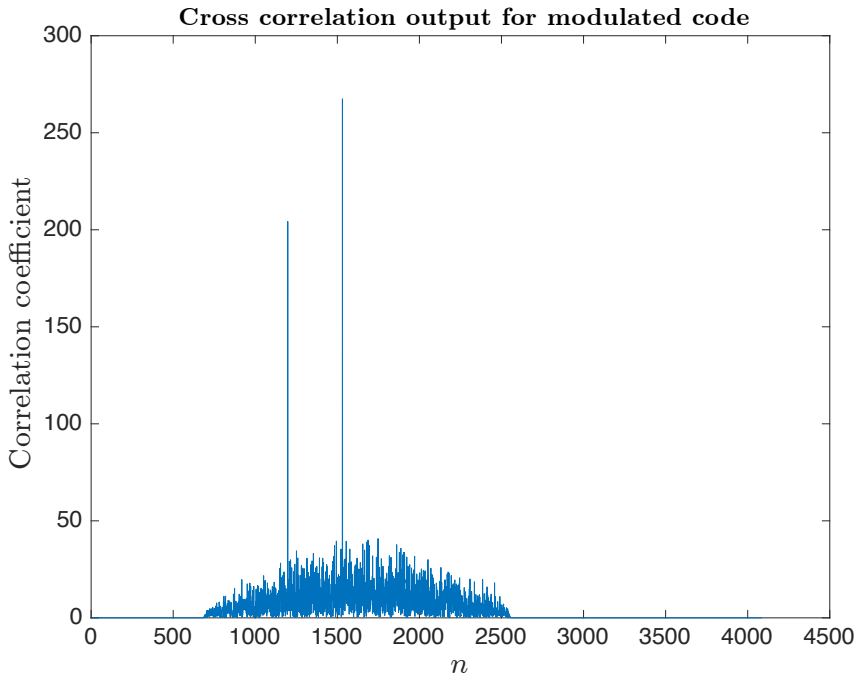
All measurements presented has been conducted with the full-rate preamble correlator implementation described in section 3.5.3.2.

An error in the implementation was discovered during result analysis. The local preamble $\dot{m}_{pre}[n]$ used in the preamble correlator was not transposed from a sequence with values [0,1] to a signal with values [-1,1]. The implication is a correlation output with more noise and spikes; thus, the experimentally determined threshold value Υ_{pre} is likely to be set higher than would have been necessary if the implementation was as intended. It is hard to estimate what this error means qualitatively for the measurements, but better SNR results should be expected for the properly implemented system, along with a better detection performance. Figure 4.1 illustrates the problem. The plot in 4.1a is a simulation of the intended correlator output, while the plot in 4.1b illustrates the actual output. Simulations were performed with the implementation of the detector class and without any noise contribution.

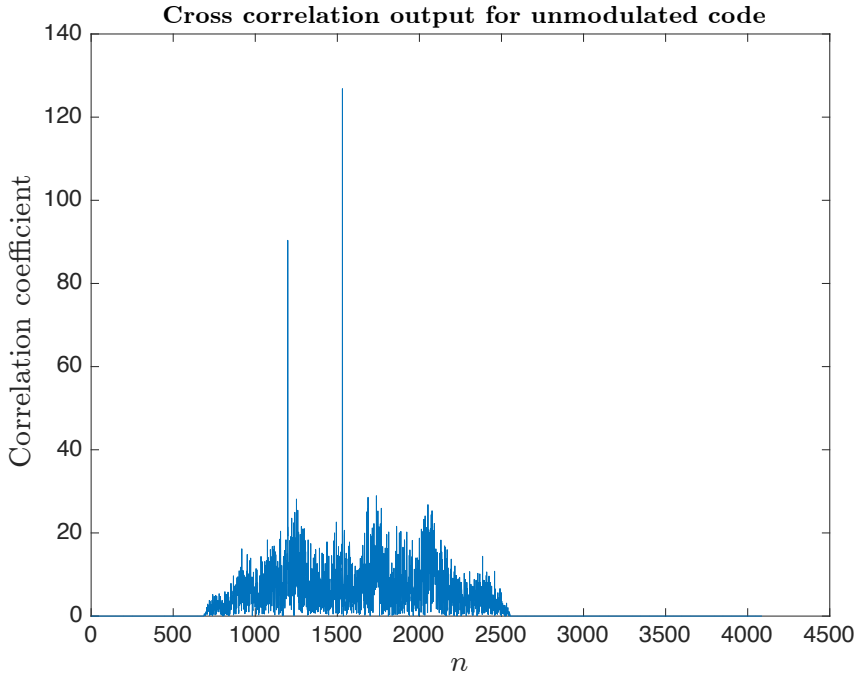
4.1 Test setup

Several measurements were conducted to analyse the functionality of the system. The majority of measurements was performed in a controlled environment in an anechoic chamber, but the system was also tested in a real environment.

The setup for the anechoic measurements was similar for all measurements. Two nodes were positioned facing each other at a measured distance of 1 meter. One



(a) Simulation of the intended correlator output.



(b) Simulation of the actual correlator output.

Figure 4.1: Illustration of simulated correlations in relation to the implementation error discovered during result analysis.

of the nodes was mounted on a turntable that was used for accurately controlling the relative angle of the node in relation to the other. Pictures of the setup can be seen in Figure B.7 in the appendix.

Standard deviation(σ) and variance(σ^2) provided in the following text uses sample difference as the presenting unit, as this represents the measured time difference between channels. Thus, the resulting values may be seen in the light of uncertainty due to system detection capabilities. For the calculated angles presented in figures, the standard deviation is presented with degrees as the unit. These deviations describe performance relevant to the total system model. When presented, a subscript will identify the interpretation, either σ_{sample} or σ_{deg}.

4.1.1 Measurement equipment

The measurements have mainly been performed with the equipment and software described in section 3.1.1. In addition, some components were utilized to verify the setup and conduct the measurements:

Bosch PLR30 Laser distance meter
Used to measure the distance from node to node in measurements.

Apple computer
Used for generating noise in the SNR measurements.
Computer type: Apple Macbook Pro 15" 2010
Processor: Intel Core i7 2.66 GHz
Memory: 8 GB 1067 MHz DDR3

RME fireface 800
Audio interface for noise transducer.

Quad 50E amplifier
Used for powering noise transducer.

A complete list of equipment with serial numbers is listed in Appendix B.1. The default measurement location was the anechoic room in the Acoustics Department at The Norwegian University of Science and Technology. In addition, measurements were conducted in an auditorium, EL5 in the Electronics Building, to investigate the real-world properties of the system.

4.1.2 Measurement uncertainty

Although measures were taken to make sure measurements was conducted as accurately as possible, all setup was adjusted by hand and confirmed by tape or laser distance meter. For positioning of the 1/4" microphones, the acoustic center of

the capsules were considered as the reference point. Deviation of ±2mm should be expected for the microphone positioning between each measurement setup.

The reference distance of 1m was calibrated at the startup of each session to compensate for the internal processing delay in the real-time MATLAB implementation. Thus, consistency for a given value of processing delay has not been assessed.

For the auditorium measurements, the distances were measured with a laser distance meter, but the deviation in distance and angle was expected to be higher at larger distances, due to the challenges of positioning the nodes along an exact line. The obtained accuracy was considered sufficient for evaluating the performance system.

4.1.3 Adaption of angle calculations

As discussed in section 2.3.1, no simple equation exists for the geometry of the extended wave travel length due to head curvature. As stated, the plane wave equation is a good approximation for angles where both receivers have line-of-sight to the transmitting node. However, when the head rotates beyond a given angle ($\approx \pm 33°$), one of the microphones will no longer obtain line-of-sight to the transmitter. This leads to extended travel length for the signal to be received by the obstructed microphone.

For the test setup, an attempt was made to compensate for this extended travel at angles exceeding line-of-sight. The correction was performed by adding a linear term to the plane wave equation for larger angles. The following procedure was used to obtain the correction:

- Acquisition of the time difference between microphones as a function of the real head rotation angle by measurements was performed.
- The maximum time difference between microphones of the test model at 90° was estimated.
- The point where the line-of-sight ends were found, such that the starting point for compensation the compensation algorithm was found.
- A linear term was calculated
- The compensated calculation was adjusted fit to test measurement

Figure 4.2 shows the difference in samples for the measurement (x-axis) plotted as a function of the head rotation angle (y-axis). The plane wave calculation(blue line), shows the angle calculated from the plane wave approximated formula(2.4), hence without the obstructing head. As expected, it results in a larger angle than the measured result. For extended travel caused by the head, the simple formula results in a complex angle value, seen as 90° for sample difference above ≈ 120. The red line shows the adapted angle calculation for compensation of the extended travel length. The angle where one of the microphones no longer have line-of-sight to the transmitter of the transmitting node is indicated in Figure 4.2 as the breakpoint.

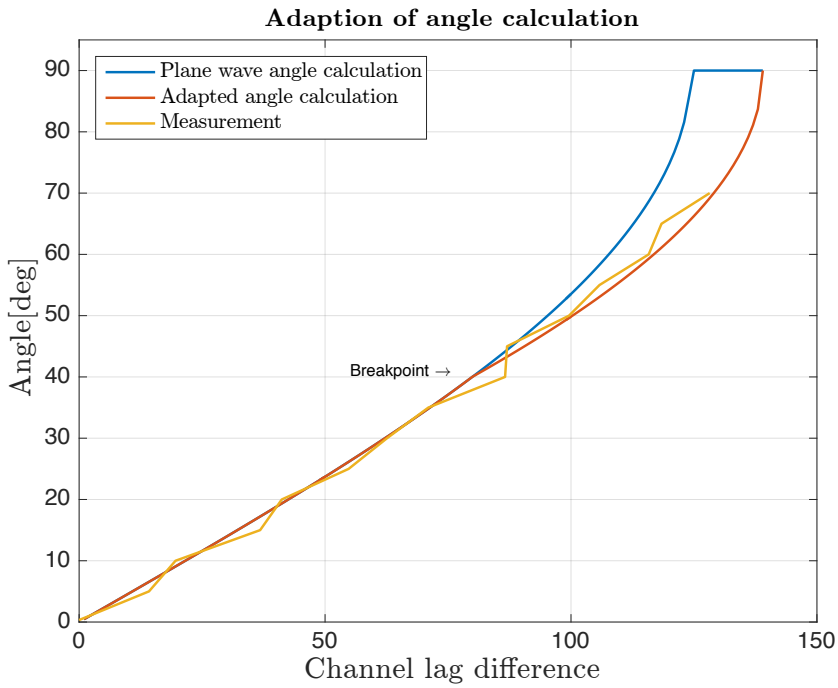


Figure 4.2: Adaption of angle calculation

The expected maximum of lag difference between channels is fitted to coincide with resulting angle of 90°. Further explanation is provided in Appendix B.1.

The figure shows that the adapted angle calculation follows the measured lag difference to a degree that was considered acceptable for the purpose, with a standard deviation for lag $\sigma = 0.4$ for angles measured $[0°, 70°]$. For angles above 70° the microphone behind the obstructing head was not able to detect the transmitted signal, leading the implementation to detect the source at 90°.

The zigzag behaviour of the measurement in the figure should also be noticed. For example, 40° and 45° incidence results in approximately the same sample difference between channels(≈ 86). This is highly likely caused by reflections from the head that was used in the setup. The issues with reflections will be discussed further in the following sections. The adapted angle calculation was only used for the default setup with microphones placed on top of each earmuff.

4.2 Anechoic measurements

To evaluate the performance of the system, several measurements was conducted in the anechoic room. A list describing the measurements and its purposes is presented below.

Attenuation caused by obstructing head
To verify the effect, the head would impose on the signal level.

Microphone position verification
A range of measurements was conducted to check if the microphone position was usable. The spacing and location were varied, and reference measurements without the head were also investigated.

Controlled reflection measurements
A range of measurements was conducted to study how reflecting objects would affect the system.

Large-quantity measurements
A range of measurements was performed at different angles to study the precision and variance of the system. Around 500 measurements were performed for each angle.

SNR measurements
Measurements with noise was conducted to study the system performance in noisy conditions.

Continous rotation measurements
Measurements with continuous movement of the head was conducted to investigate the performance while the head is rotating. The purpose is to investigate the system performance and signal detection abilities for a rotating head.

IMU verification measurement
Measurements were conducted to investigate the IMU performance.

4.2.1 Attenuation caused by obstructing head

To verify the prediction of attenuation caused by the head shape, the sound pressure level was measured at different angles. The measurement was performed in 1° increments for angles 0° to 90°, excited with a stationary 40kHz sine signal. The default microphone positions were used, and one microphone was considered.

The measured attenuation relative to the on-axis response can be seen in Figure 4.3. For the first °, some dips and peaks in the curve are seen. This is due to constructive or destructive interference from the head or earmuffs. As expected, the trend of the curve is increasing attenuation for increasing angles. At around 25°, the pressure begins to drop more rapidly, as the limit of line-of-sight is approaching. For angles larger than $\approx 75°$ the response is unpredictable, and indicates that the noise floor has been reached.

4.2.2 Microphone position verification

To verify that the selected microphone positions will provide the necessary accuracy, a number of test were conducted. As mentioned in section 2.5, the default

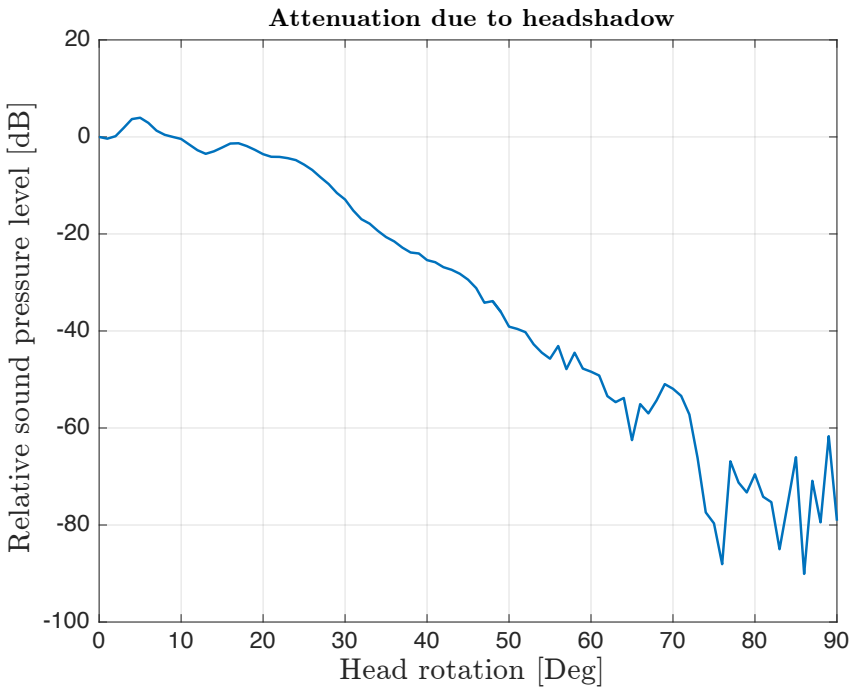


Figure 4.3: Relative attenuation caused by headshadow for increasing angle.

microphone position is on the top of each side of the earmuffs with a spacing $d = 220$mm, thus simulating that they are an integrated part of the earmuffs. This position is subject to reflections from the head.

Other configurations are assumed to provide better results, in particular, measurements where the microphones were placed above the head instead. Placing the microphones above the head would omit the difficulties with the extended travel length imposed by the obstructing head, as well as the expected loss in signal strength leading to lower detection rate for the correlator. This configuration may also provide information about the precision of the detection system.

Three configurations with the microphones located above the head various spacings were tested. The default configuration without the head was also tested to investigate the impact of reflections caused by the head. Figure 4.4 illustrates the different configurations. In all configurations, the ultrasonic transmitter and the microphones were aligned with the axis of the rotating head stand. The height of the microphones is 10mm above the hearing protector headband to the acoustic centre of the microphones. All measurements were performed with ten measurements per angle, and with an angle incrementation of 5°. The measurements with the default setup will be discussed in section 4.2.3.

Figure 4.5 presents the results for the measurements with the microphones located above the head with 50 mm spacing. The measurements show consistent results,

50 mm spacing top

130 mm spacing top

220 mm spacing top

220 mm spacing default setup

Microphone

Figure 4.4: Microphone position verification setup

considering the low deviation from the mean, maximum $\sigma_{deg} = 2.49°$ for the 60° measurements.

The resulting angle calculation is not consistently accurate. Most of the error are probably caused by reflections from the head. Examining the raw data of the measurements at 55° indicates a travel length of 1.3cm longer than should be expected at this angle. Measurements of possible reflection trajectories of the physical model correspond well with a reflection caused by the forehead of the model.

The reduced resolution due to the shorter spacing of microphones will naturally contribute to reduced accuracy. Using equation 3.2, the maximum angular resolution is calculated to be 4° for on-axis measurements.

For the measurements with 130mm spacing, a test sequence was performed both with and without the head to verify the effect the head impose as a reflecting surface. Figure 4.6 shows the results with and without the head. The figure shows strong indications of the negative effect caused by head reflections.

For the no-head measurements the figure also gives an indication of the reduced accuracy for increasing angle as discussed in section 3.2, given the increasing standard deviation for increasing with angle[1]. Maximum standard deviation for all angles in sample difference was $\sigma_{sample} = 0.527$, equal for both measurements.

[1]Figure B.3 in the appendix show the same figure plotted as resulting deviation in samples,

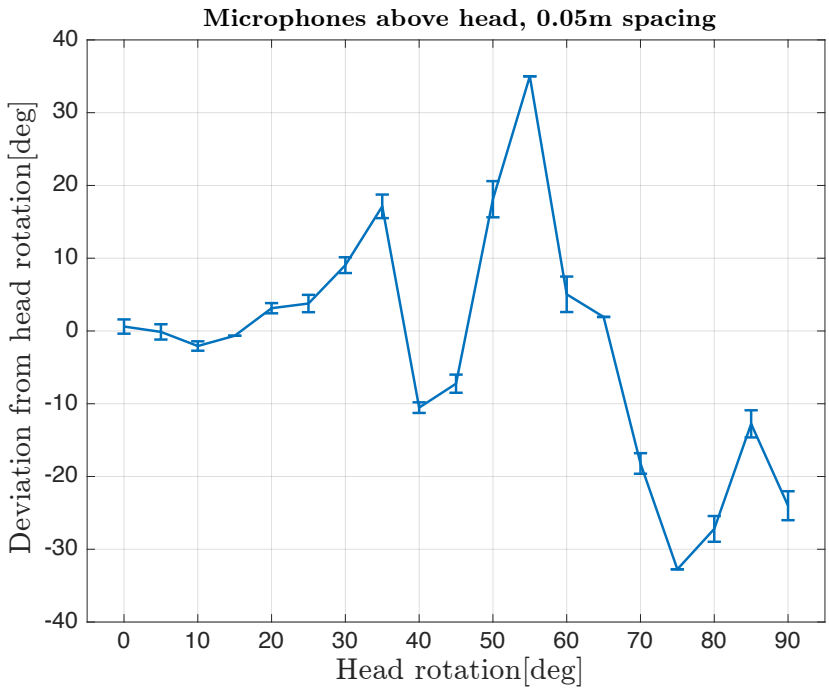


Figure 4.5: Deviation between calculated and physical angle for 50mm microphone spacing, 10 measurements per angle. The data points show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

Figure 4.7 shows result from measurements performed with the microphones placed above the head with the same spacing as the default setup(220mm), with and without the head present. The results suggest that the increased spacing reduces the influence of reflections leading to errors in angle calculation compared to the 130mm setup. The maximum sample standard deviation for the measurements was $\sigma_{sample} = 0.527$, equal for both measurements.

4.2.3 Large-quantity measurements

To evaluate the precision in a statistical perspective, a test sequence of 500 measurements for every 10° increment was performed for angles between 0° and 80°. Figure 4.8 shows the resulting calculated angles for the 0° to 70° measurements. For the 80° measurements the channel obstructed by the head was not able to correlate the signal leading to a calculated angle of 90°. The angle correction described in Section 4.1.3 is used.

As can be seen in the figure, there is an ambiguous behaviour for the calculated

indicating more similar standard deviation for increasing angle

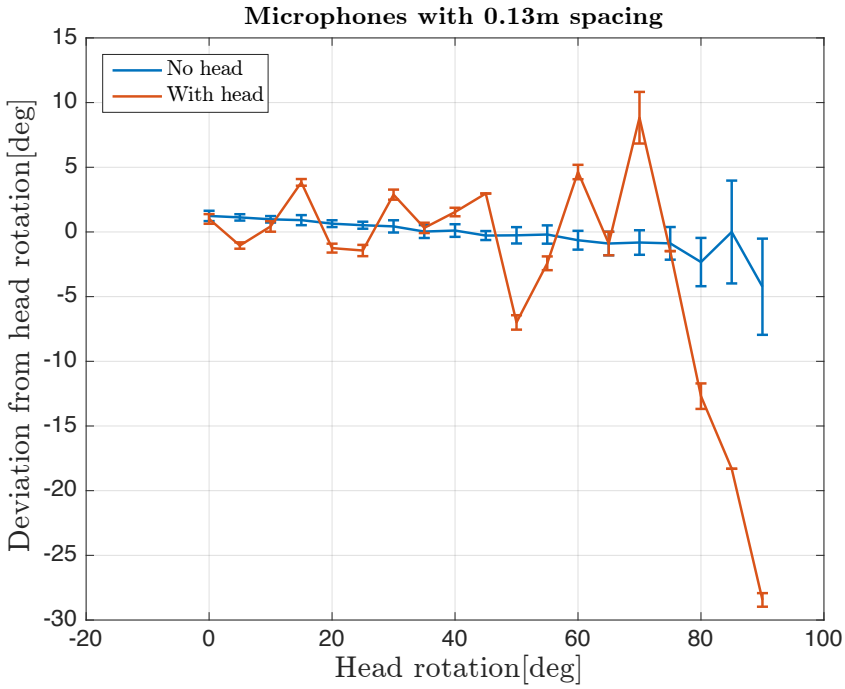


Figure 4.6: Deviation between calculated and physical angle for 50mm microphone spacing, 10 measurements per angle. The data points show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

angle in relation to the rotation angle. For the smallest angles[0°20°], the deviation increases with increasing angle, indicating a longer travel distance than the rotation of the head should imply. For angles above 30°, the tendency of the calculated angle is below the real angle of the rotation. This indicates that the difference in travel length between channels is shorter than the rotation angle should require. The 60° angle is accurate. This stochastic tendency of acquired angle was seen in other measurements as well (Figure 4.6) and is plausibly caused by head reflections. The standard deviation and variance between measurements is however low, $\sigma_{sample} < 0.49$ and $\sigma^2_{sample} < 0.38$ for all head angles. This indicates that the system is stable in regards to precision for multiple measurements. The variance and standard deviation of all angles can be seen in the appendix, table B.3.

Figure 4.9 shows the calculated distance for the large-quantity measurements. Increasing calculated distance is observed for an increasing angle. The reason for this is in the implementation. When both channels correlate the incoming signal, the system returns the mean of the two calculated lengths, and due to the extended travel length caused by the obstructing head, the mean will be slightly higher for steeper angles.

Some obvious errors in the distance calculation have been omitted in the figure for

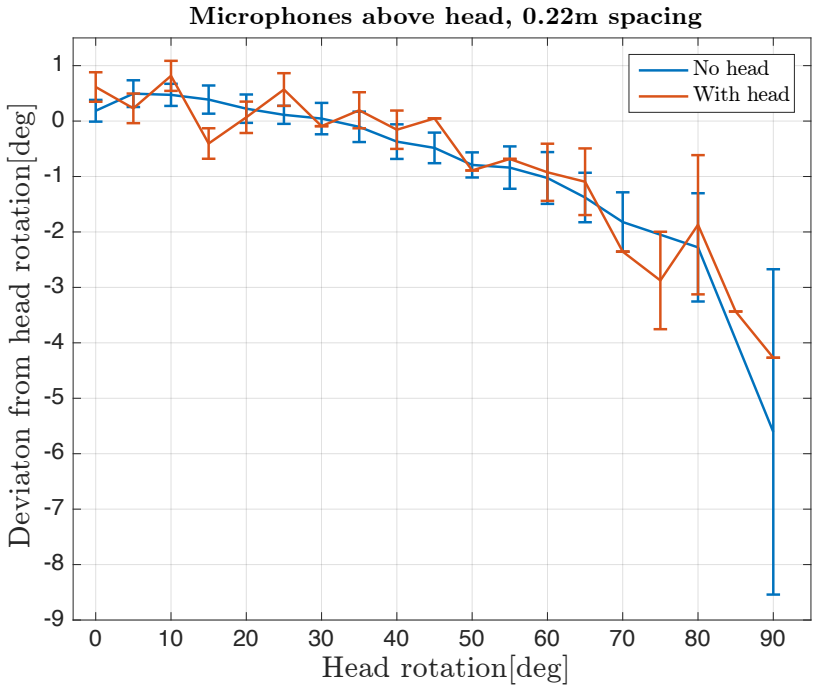


Figure 4.7: Deviation between calculated and physical angle with 0.22m spacing with and without head. Datapoints show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

better viewing[2]. The obvious errors that were returned in seven of the measurements were in the magnitude of ±23m and ±300m, and was caused by a bug in the implementation. The bug is probably due to absolute time miscalculation caused by a logic error. Due to the bug, the probability of error is $P(error) = 0.0017$ for the distance measurement.

Omitting the distances caused by the bug, the resulting distance show very little deviation for each angle with maximum encountered $\sigma_{distance} < 0.63 \cdot 10^{-3}$m and $\sigma^2_{distance} < 0.39 \cdot 10^{-6}$m. Standard deviation and variance for respective angles can be seen in the appendix, Table B.4.

4.2.3.1 Controlled reflection measurements

A series of measurements was performed with reflecting surfaces in the anechoic room. The goal was to inspect how the system performs with discretely quantified reflections that could be expected in a real life environment and investigate the implications of reflection for the detection system.

[2]The full dataset can be seen in the appendix, Figure B.5

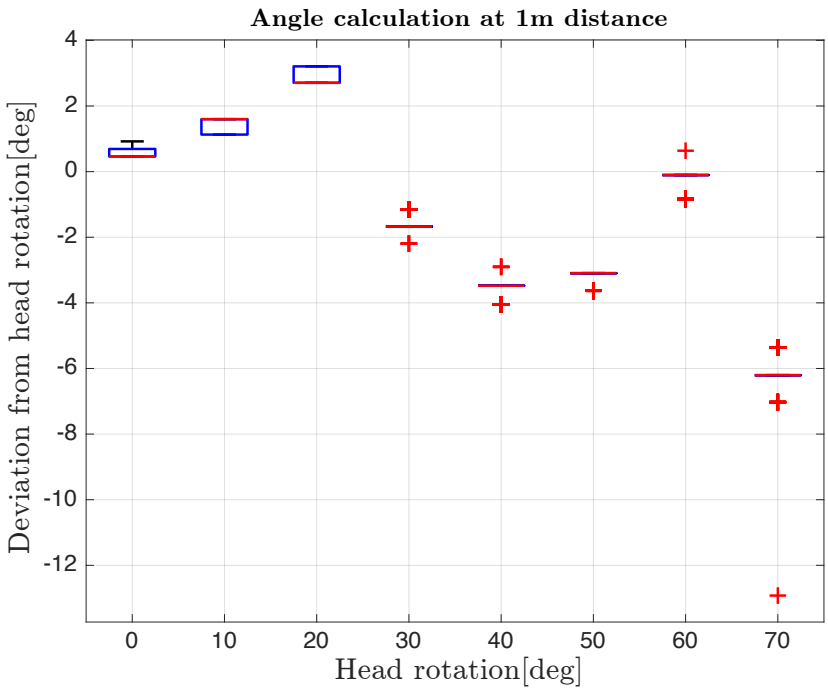


Figure 4.8: Deviation between calculated angle and real angle of rotating head. 500 measurements performed for each angle[a]. For angles above 30°, the boxes representing the distribution are to small to be viewed, indicating small deviation.

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the data points that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + $1.5 \cdot$ boxrange, or lower than 25% percentile- $1.5 \cdot$ boxrange.

The different setups that were tested can be seen in Figure 4.10.

Figure 4.11 shows the results from measurements with side and back wall, denoted setup a) and b) in Figure 4.10. The two cases both show stochastic behaviour around the correct angle. The zigzag trend of the resulting angle is most likely caused by head reflections, as the differences in travel length would impose angles outside the physical definition. A reflection from the side wall would by simple algebra result in a extended travel length of $\approx$ 0.4m. This can be verified by inspecting the measured impulse response of the situation. Figure 4.12 shows the correlator output, and the reflection can be seen lagging $\approx$ 0.4m behind the direct signal.

For setup b), the reflections from the back wall were not detected, partly due to the characteristics of the microphones. Figure 4.13 shows the impulse response of the received signal at 80°. The reflected signal from the back wall can be seen lagging 0.85m behind the first correlation, which corresponds well with the distances in

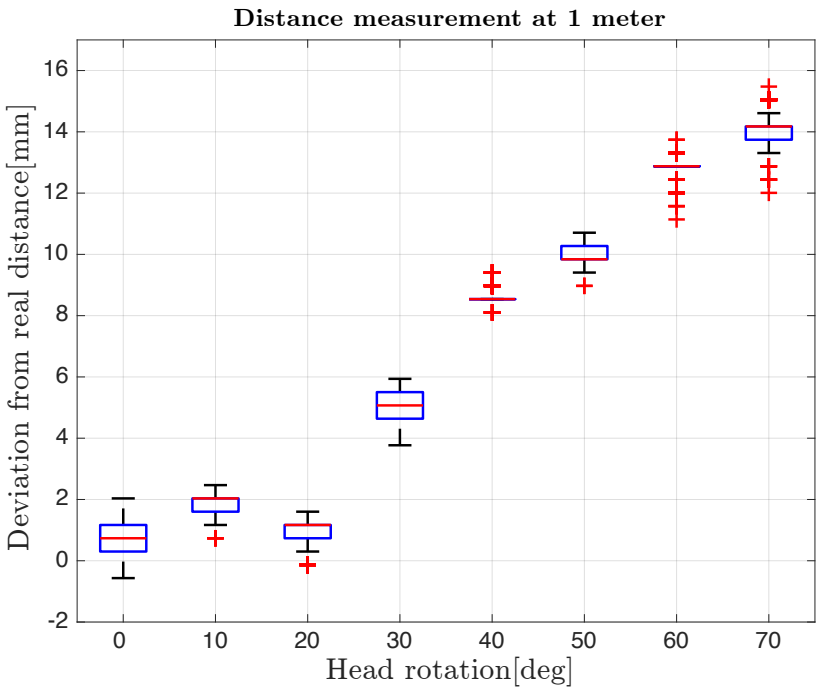


Figure 4.9: Distance measurements for increasing angles with 7 obvious errors omitted for better viewing. 500 measurements performed for each angle. For 40° and 70°, the boxes[a] representing the distribution are to small to be viewed, indicating small deviation. Notice that Y-axis denoted in mm.

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the data points that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

the setup. The reflected signal from the head can also be seen. The distance lag corresponds well with the physical measured distance between the microphone and head, 45mm.

As the head turned away from the wall for the side wall measurements, the directivity of the microphones reduces the signal level of the reflected sound wave. (Consider Figure 3.2 for directivity pattern). The results for angles above 75° are omitted in the figure since the system was unable to successfully detect the signal on both the left and right channel.

For the slanted wall, situation c) in Figure 4.10, the angle deviation is presented in Figure 4.14. As opposed to situation b), this measurement is designed to test reflections from a facing surface. Thus, the directivity of the microphones will not contribute to reduced reflection effects. Clearly, ambiguities were introduced by the reflecting wall. Figure 4.15 shows the sound wave paths that caused the erroneous

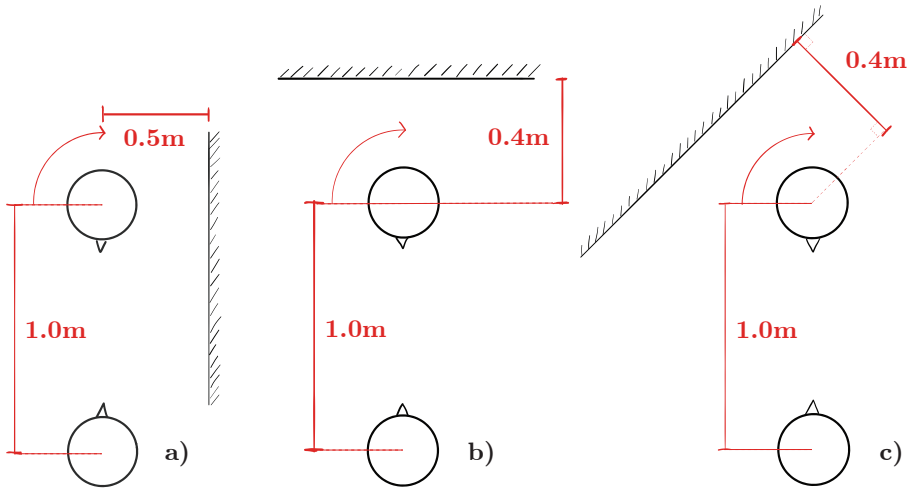


Figure 4.10: Different setups for measurements with reflecting surfaces.

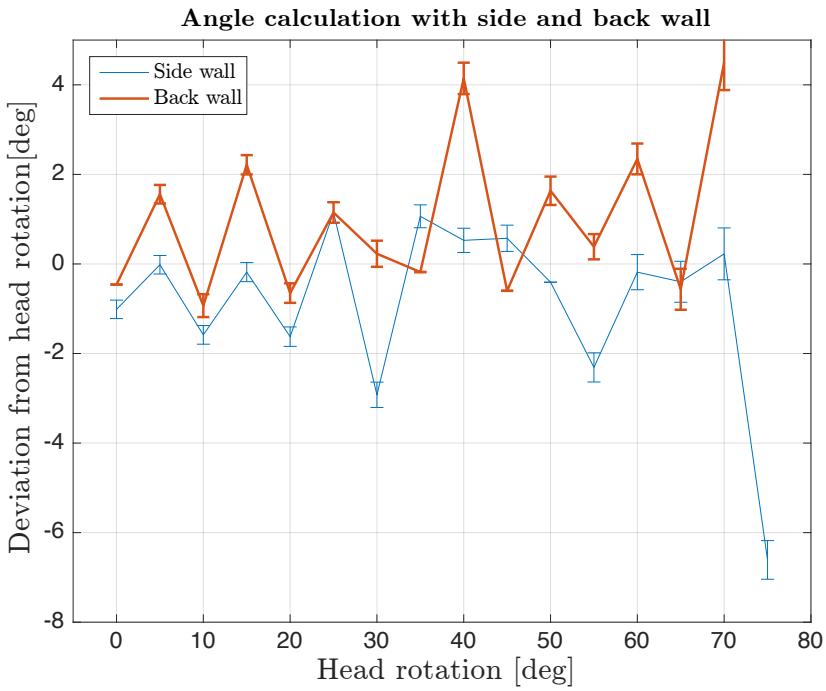


Figure 4.11: Angle measurements with side and back reflection. Datapoints show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

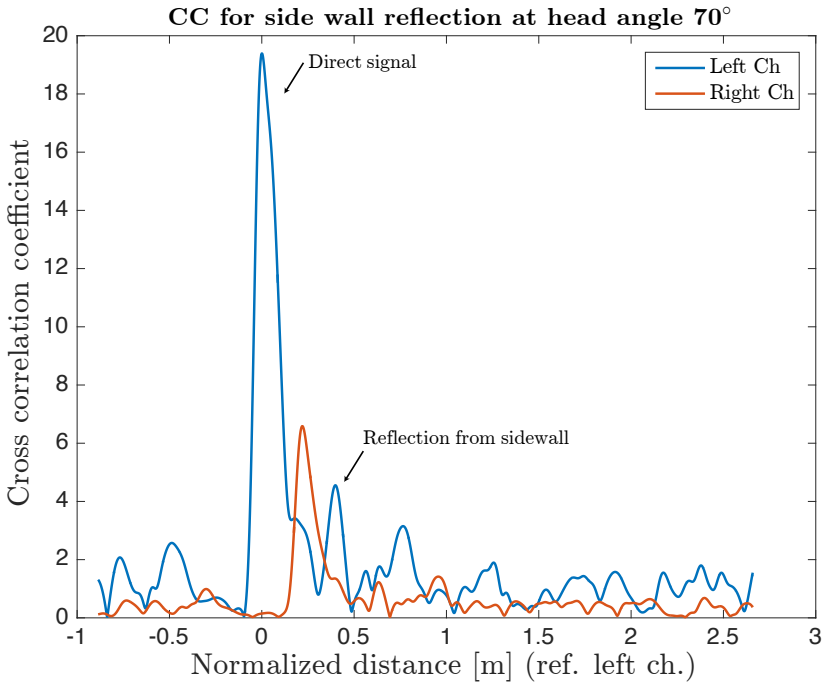


Figure 4.12: Impulse response of the received signal at angle 70°. The reflection from the sidewall can be clearly seen ≈ 0.4m behind the first correlation, corresponding with the extended travel from side wall reflection.

results. Figure 4.16 shows the measured TDOA for each angle, but recalculated to present the travel distance difference. The expected (correct) angle is included for reference.

At a given angle(≈ 45°), the reflected signal yields a higher correlation coefficient than the direct signal for the channel obstructed by the head and is thereby selected for calculation. Figure 4.17 shows the impulse response of the situation, denoted 1) in the figures. At larger angles(situation 2) at ≈ 70°, the reflected signal is stronger than the direct for both channels, which in turn leads to a shorter time difference between channels thus leading to calculated angle in the wrong quadrant.

Figure 4.18 shows the impulse response of the received signals, showing the left channel lagging the right channel for the highest correlation coefficients. The directivity of the microphone will influence this type of behaviour, considering the response is approximately 30dB lower at 90° than on-axis for the microphones used, discussed in section 3.1.1.5, as well as the attenuation caused by the head shadow.

Both situations may be prevented in the implementation;

1) By implementing a maximum allowable difference between channels, ignoring signals on the corresponding channel arriving to late.

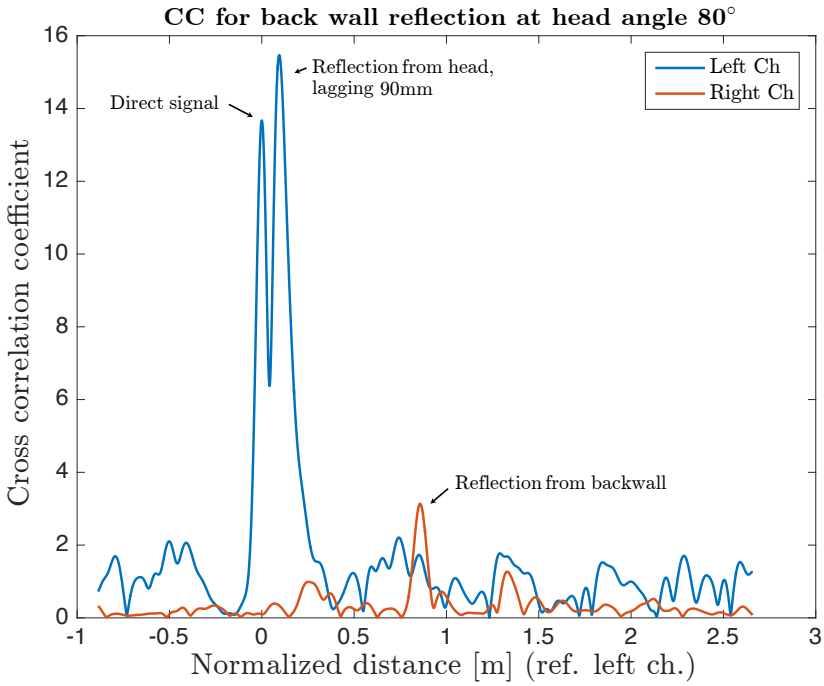


Figure 4.13: Impulse response of the received signal at angle 80°. The reflection from the back wall can be seen lagging 0.85m behind the first correlation. Reflection of the head can be seen lagging 0.09m behind the direct signal.

2) By using the first incident sound as reference related to the reflected signals, instead of the highest correlation coefficient.

For very large angles, the direct sound attenuation may be problematic, and situations like 2) might still be a problem. However, selecting microphones with omnidirectional characteristics may help.

4.2.4 SNR performance

An important performance aspect is how the system perform in noisy conditions, as well as the sensitivity to near-far effects. Therefore, signal-to-noise ratio(SNR) measurements were performed. The measurements were performed in anechoic conditions, with the nodes placed on-axis at 1m. To generate the noise signal, the same type of transmitter used for the nodes was utilised. The transmitter was placed right next to the node transmitter on the responding node, to ensure that the characteristics of the signals are the same.

The noise signal used was other randomized gold codes summed together in randomized order. This leads to a noise signal that is identical to the signal in spectral

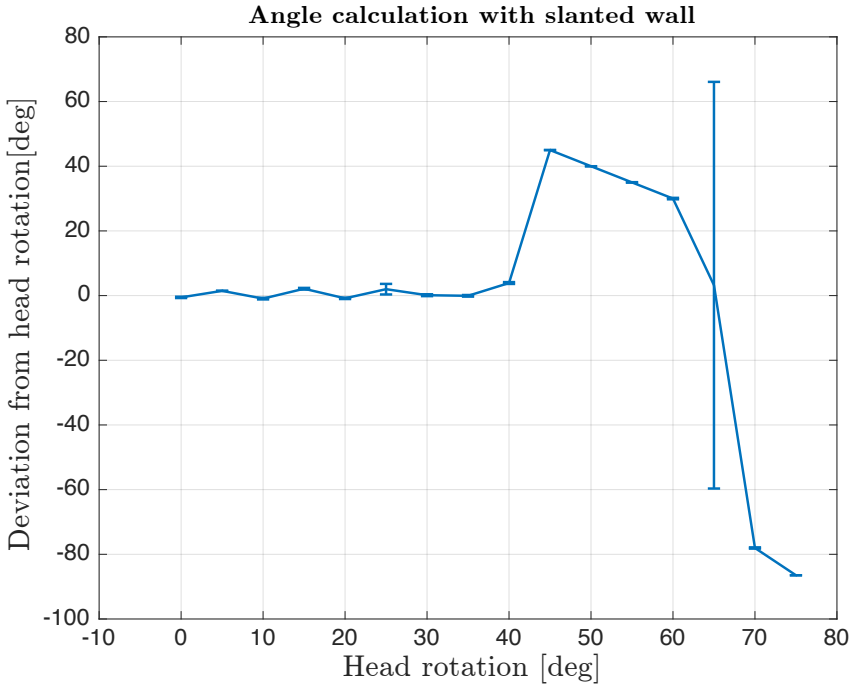


Figure 4.14: Angle measurements with reflections caused by slanted wall. Data points show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

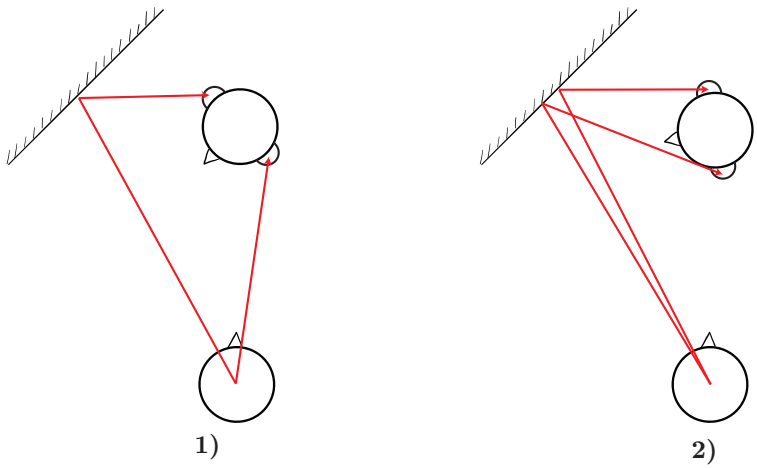


Figure 4.15: Reflection paths for slanted wall.

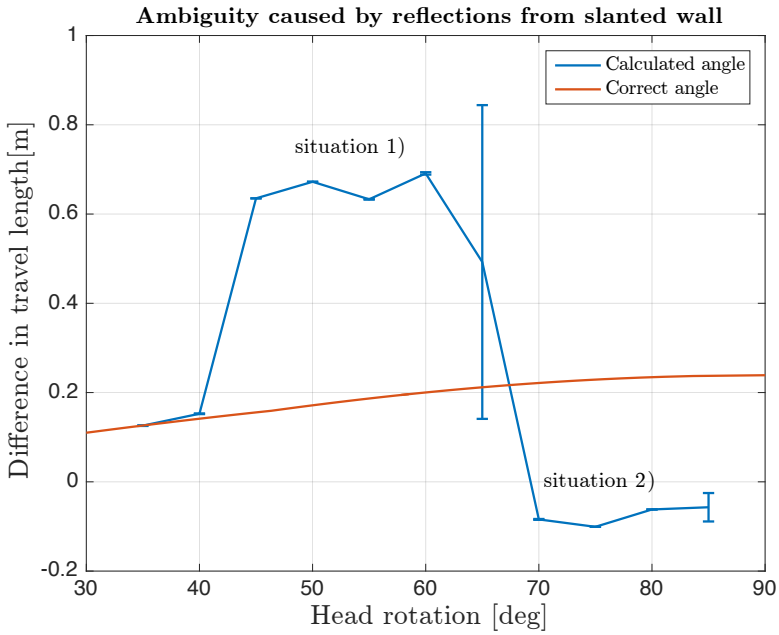


Figure 4.16: Ambiguity caused by reflections from slanted wall.

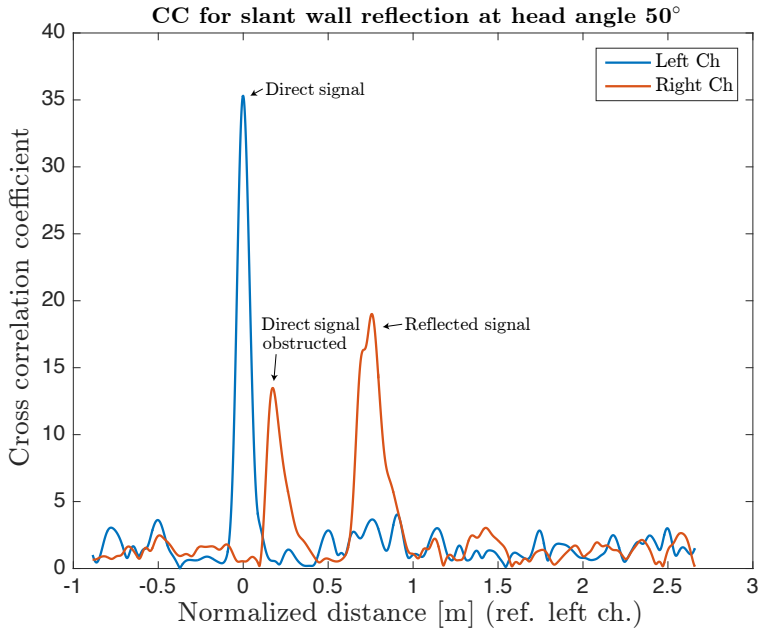


Figure 4.17: Impulse response for situation 1). The reflected signal is stronger correlated than the direct signal.

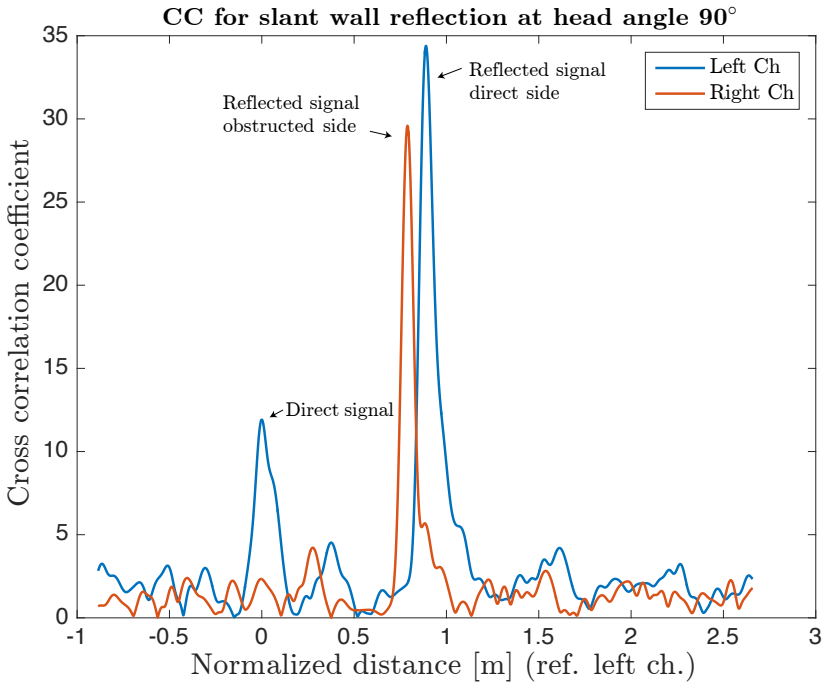


Figure 4.18: Impulse response for situation 2). The reflected signal is stronger correlated than the direct signal for both channels thus leading to wrong angle calcuation.

characteristics. The noise signal was generated with a standalone computer running MATLAB. The noise level was then calibrated by measuring and comparing the rms values of both the signal and noise. Between 450 and 695 measurements was performed for each SNR level in the range of 0 to -6dB.

Table 4.1 shows the probability of detection error for the different signal-to-noise ratios. In the table, the detection error is defined for two different situations. The first case is when only one channel is able to detect the transmitted signal, thus being able to calculate the distance to the other node, but not the angle. In the other case, none of the channels are able to correlate the input, resulting in no detection. The results show a clear threshold for where the system no longer is capable of correctly detecting the transmitted signal due to noise. Due to the implementation error described in the introduction in this chapter, the SNR performance is probably better than the results presented in this section. Comparing to the results obtained in the simulation in Section 3.5.4, the real performance obtained is significantly better. However, the effect of the transducer and channel is not considered in the simulation.

Table 4.1: Detection error of measurements performed at various SNR

SNR[dB]	**0**	**-1**	**-2**	**-3**	**-4**	**-5**	**-6**
Probability of error 1 ch.	0	0	0	0	0.004	0.556	0.970
Probability of error both ch.	0	0	0	0	0	0.007	0.221
Measurements performed	522	517	695	608	574	450	521

4.2.5 Continuous rotation measurements

In a real-life situation, the system will need to perform with the nodes moving in arbitrary directions. Doppler shift is not assessed in this work, and it is expected that the system will be limited to detection of static positions. To investigate the performance of the system in motion, measurements involving continuous rotation of the head was performed. Linear movement has not been tested due to lack of equipment for the purpose. The test is limited to investigate whether the system is able to successfully correlate the transmitted signal and calculate the angle and distance for a given angular speed. The actual physical angle at the moment of transmission is not recorded as this would require deeper integration of the turntable control code into the positioning system. Thus, the accuracy of the measurement in relation to the real angle is not verified. The result of this test merely serves as a verification of detection capabilities for a rotating head.

Two series of measurements with different rotational speeds was performed with head panning $\pm 60^\circ$ in the horizontal plane. Two different rotational speeds was tested, denoted fast and slow, with respectable angular velocity $\omega_{fast} = 36.1\,\mathrm{deg/s}$ and $\omega_{slow} = 7.2\,\mathrm{deg/s}$. 500 measurements were performed for each speed at 1m distance. It should be noted that due to deceleration and acceleration of the angular speed in the boundary angles, the distribution of angles will not be perfectly uniform at the boundaries.

Figures 4.19 and 4.20 shows the distribution of the acquired calculated angles for the two different speeds. When examining the figures, one may note that the measurements with slow speed results in a more even distribution, whereas for the fast rotation the distribution is far more uneven. The cause of this distribution has not been further investigated, but systematic errors as time synchronization due to periodic transmission times coinciding with angular speed may be a cause.

The length of the signal used in the implementation is assumed to affect the accuracy of these results for the higher angular velocity. Given the time needed to transmit the pulse, $t_{transmission} \approx 160\mathrm{ms}$ (as discussed in section 3.2), the head will rotate $\approx 5.8^\circ$ during the correlation process in the high-speed measurement.

In the high-speed case, 78 of the 500 measurements($\approx 15.6\%$) was not detected in both channels. These measurements are presented as the red bars in Figure 4.20. However, in the low-speed case, all measurements was detected at both channels.

The distance measurements for the two series can be seen in Figure 4.21. In the high-speed series, there were six measurements calculated incorrectly, omitted in the figure for better presentation. The total figure including erroneous results

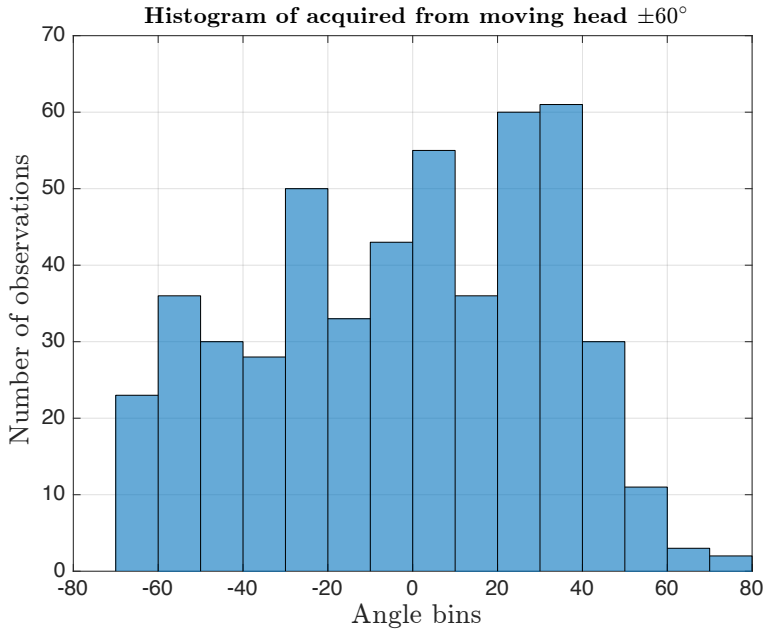


Figure 4.19: Distribution of calculated angles for head rotating ±60°, with angular velocity $\omega_{slow} = 7.2\,\mathrm{deg/s}$. 10° in each bin.

Table 4.2: Standard deviation and variance for distance measurements performed with continuously rotating head.

Angular velocity[deg /s]	$\omega_{slow} = 7.2°/\mathrm{s}$	$\omega_{fast} = 36.2°/\mathrm{s}$
σ [mm]	4.1	7.1
σ^2 [mm]	0.017	0.050
Measurements performed	500	500

can be viewed in the appendix, Figure B.2. The cause of the erroneous distances is probably due to the earlier mentioned implementation bug, as the resulting distances are in the same range. Figure 4.21 verifies that the system was able to calculate the distance fairly accurate even though the angle could not be calculated for the high-speed case. Standard deviation and variance for distance calculations given in Table 4.2.

From these results, one may conclude that the Doppler-shift causes detection failure, even at a very low displacement of the node during signal transmission/reception. However, for successful detections, the results are precise.

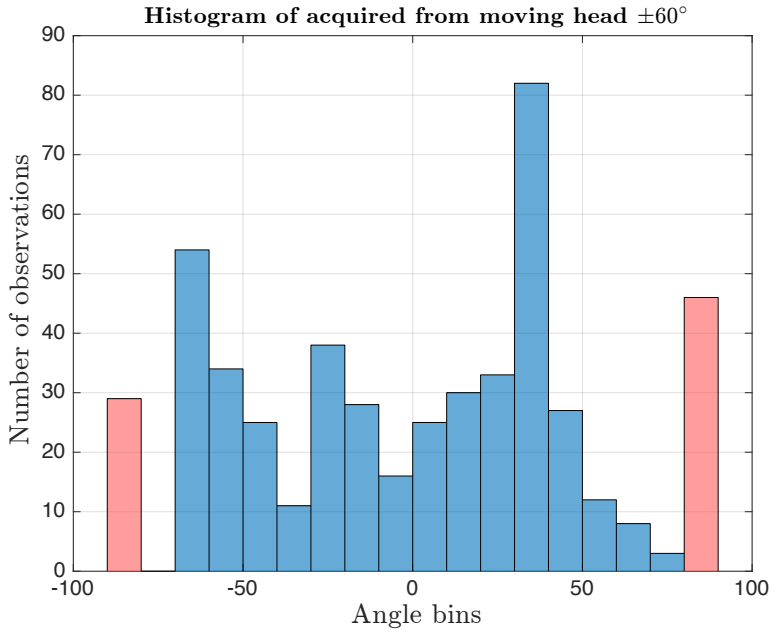


Figure 4.20: Distribution of calculated angles for head rotating ±60°, with angular velocity $\omega_{fast} = 36.1\,\mathrm{deg/s}$

4.2.6 IMU verification measurement

A test measurement was conducted to verify the functionality of the IMU implementation. As explained in section 2.3.4, the task of the IMU is to keep track of the head rotation of the node, and thus update the position map for all other nodes faster than the positioning system itself is capable of. The tests verified that the simple IMU implementation was able to measure the rotation with good precision. Figure B.4 in the appendix shows the difference between the physical head rotation and the output from the IMU. The measurements showed that the deviation increased from 0° to −0.4° during a measurement series that lasted for approximately 5 minutes. The increase in deviation with angle may as such be a result of the lacking time invariance of the measurement. However, the precision may be considered to be sufficient.

As mentioned in Section 3.1.1.7, the implementation of the IMU was performed with the use of an iPhone 6 running the MATLAB mobile app, communicating with the real-time MATLAB implementation over Wi-Fi. The update frequency of the IMU system was in the magnitude of $f_{update} \approx 2$Hz, which was sufficient for the updating of angle in the test. These measurements were primarily performed as a proof-of-concept. Therefore, no further performance investigations was performed.

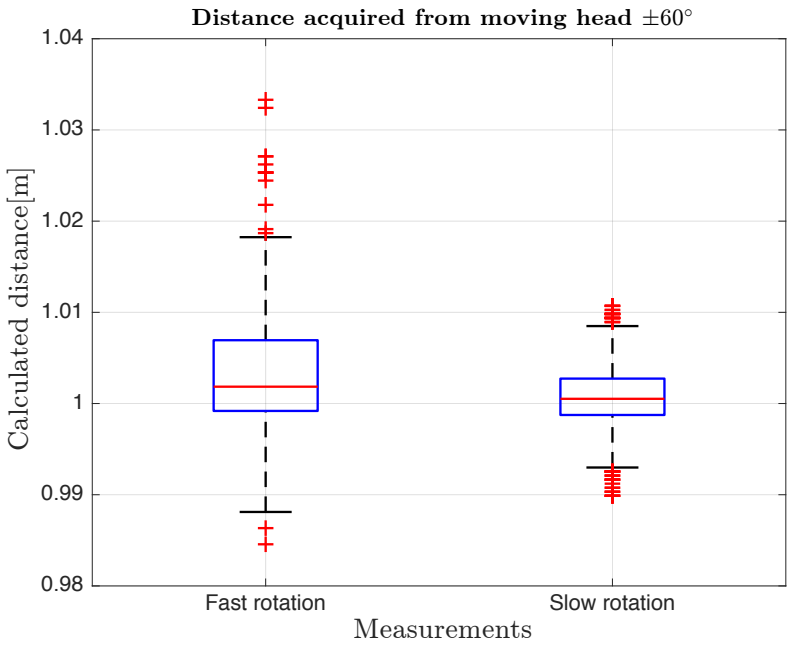


Figure 4.21: Distance calculation for head panning(±60°) in the horizontal plane, omitting far outliers[a].

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

4.3 Auditorium measurements

In addition to the controlled environment measurements, a real-world test was performed in a large auditorium. The auditorium introduces arbitrary reflections from walls, floor and nearby obstacles as expected in a real use case. The measurements includes distances ranging from 1 to 12m and angle calculation in the range of ±90°. In addition, a maximum length test was performed at 20m. An illustration of the test setup, including positions, is included in Figure 4.22. As may be seen from the figure, the static, rotating node is placed near a wall and rigid tables in one end of the auditorium. For negative rotation angles, the node points towards a hard wall spaced approx. 1.5m away, and a wall with a blackboard, distance approx 3m normal to the line between the two nodes. Pictures of the setup are included in the appendix, Figure B.9 and B.10.

The rotating node was placed in a fixed position while the distance node was moved to increasing distances. The distances tested was {1, 2, 4, 6, 8 ,10, 12, 20}. Distance [1-12] was performed along a straight line on the floor of the auditorium, while the 20-meter sequence was performed with the nodes positioned diagonally

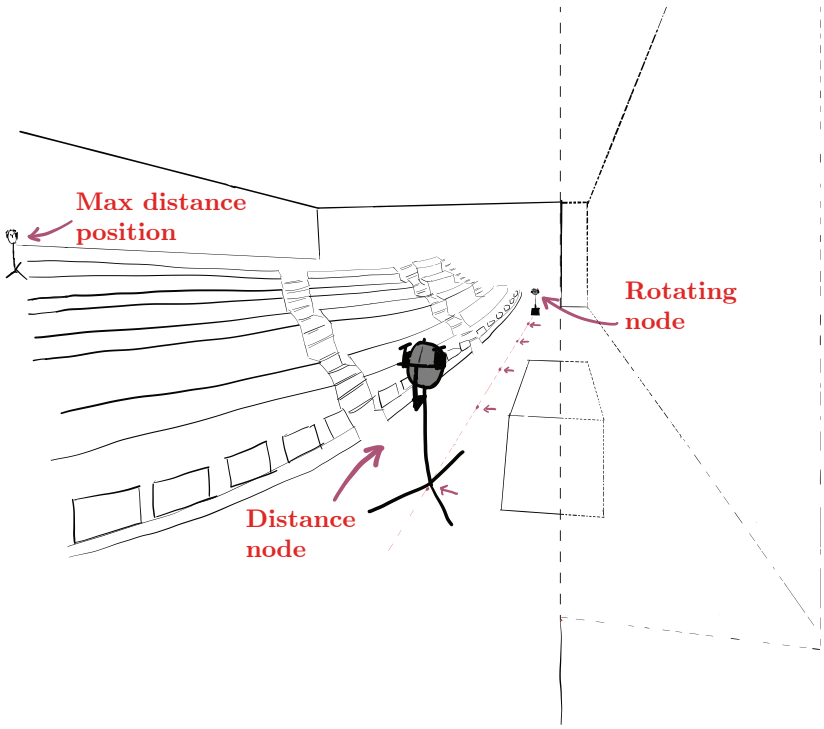


Figure 4.22: Measurement setup in auditorium

in the auditorium with an elevation angle of $\approx 6°$.

Angular precision

Figure 4.23 shows the resulting calculated angles for head rotation $\pm 50°$ and distance 1 to 8 meters. The plot is based on ten measurements per angle, for each distance. The angle calculation for distances above 8 meters did not provide sufficient data in the auditorium measurements and is therefore left out in the analysis of the results. Detection is achieved, but the time difference between the channels are greater than expected. Thus, reflections affect the measurements to a large extent, and the simple detection algorithm based on maximum correlation coefficient is not sufficient.

The small standard deviation, plotted at each point, indicates that the data is consistent within each measurement. Some erroneous trends can be seen in the results. All measurements at 1 meter distance tend to indicate smaller calculated angle than the real head angle[3]. Impulse responses was not captured for these distances. Thus, it can not be verified that the erroneous calculations for these angles was caused by reflections from the surface of the head. However, measurements of the possible trajectory indicate that it is plausible that reflections from the forehead would result in a deviation corresponding to the results in the figure.

[3]In appendix, Figure B.6 shows the calculated angles for distances 1 to 8 meters for head rotation angles $-40°$ to $30°$ for better readability regarding deviations.

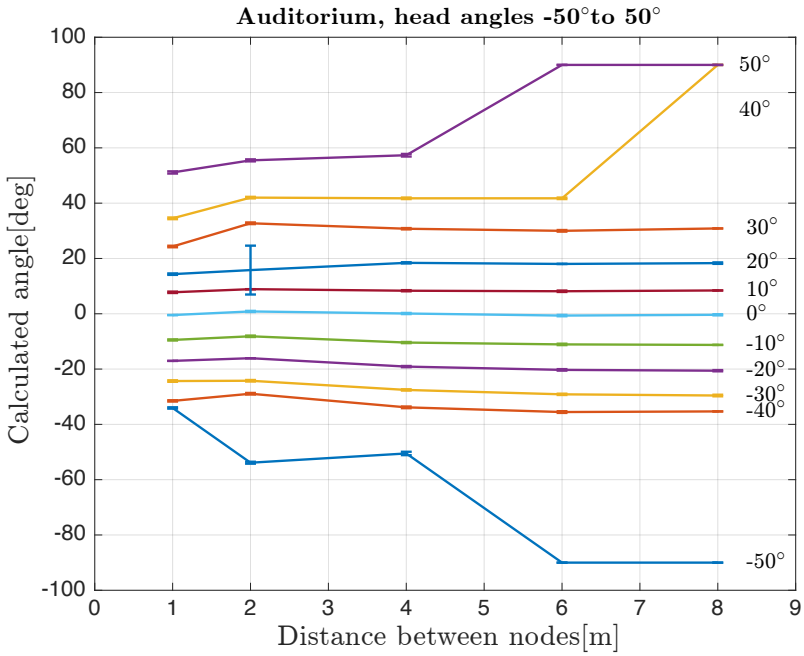


Figure 4.23: Calculated angles in auditorium for distances 1 to 8m. Each graph represents the respective real angle of the rotating head. Datapoints show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point.

At 6m distance, $-50°$ incident angle, the system consistently calculates the wrong angle. By examination of the data, the travel difference between channels was $\approx$ 2.5m, which coincides exactly with the reflecting path of the side wall. This exemplifies that the combination of head attenuation and reflective surface results in erroneous calculation. A similar investigation for the 40° calculation at 8 meter distance showed that the correlator was able to correlate both left and right channel. However, the travel difference between channels was measured at $\approx$ 0.94m ($\sigma = 9.8 \cdot 10^{-4}$). This can be explained by reflecting surfaces in the auditorium. The 50° produced similar result with a travel difference of 0.97m, ($\sigma = 8.0 \cdot 10^{-4}$). This indicates that the reflections for these angles was caused by the same reflecting surface.

Another example is investigated in Figure 4.24. The figure shows errors occurring due to reflection for neighbouring angles. The travel difference between channels should be $\approx$ 0.28m, but the measurements show that the difference between channels has quite different values. The fact that the difference is small between the two angles indicates that the signals follow the same trajectory[4].

[4]Except for at 4 meters where the standard deviation for the 80° signal is significantly larger than the other, indicating that at least two different reflection paths yields approximately the same correlation coefficient.

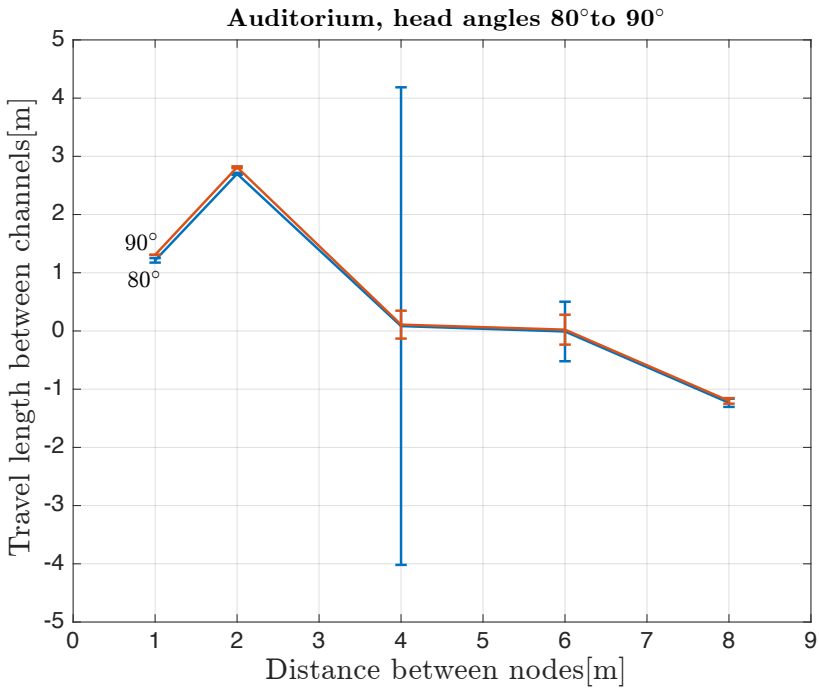


Figure 4.24: Angle calculation for head angle 80° and 90° in auditorium. The similarity indicates the same erroneous trajectory path for both angle calculation. Y-axis represents travel length[m].

Distance precision

With regards to the distance parameter, some deviation between measurements was registered. Figure 4.25, shows the results for all angles at each distance. We see in the figure that the distribution is uneven. The mean of measurements indicate longer distance than the real distance, and the deviation parameters are large ($\sigma_{all} = 0.59$, $\sigma^2_{all} = 0.38$).

Figure 4.26 presents the distance parameter for $\pm 40°$, and indicates that the deviation of distance is smaller for these angles($(\sigma_{\pm 40} = 0.22$, $\sigma^2_{\pm 40} = 0.05)$. Consistent distance is obtained up to eight meters.

The cause of the deviation can be attributed the implementation. As the distance is calculated as the mean of both channels, the microphone obstructed by the head will correlate with a signal that has been reflected off of a nearby surface for large angles $> \pm 40°$. Directivity of the microphones also contribute. Thus, the selected signal path for the obstructed microphone channel is a reflection, and a longer travel distance results.

As long as both microphone channels are able to correlate the transmitted direct signal, the deviation at the measured distances is low. Figure 4.28 shows the deviation from real distance for angles $\pm 40°$, 1 to 8m. Note that the figure presents

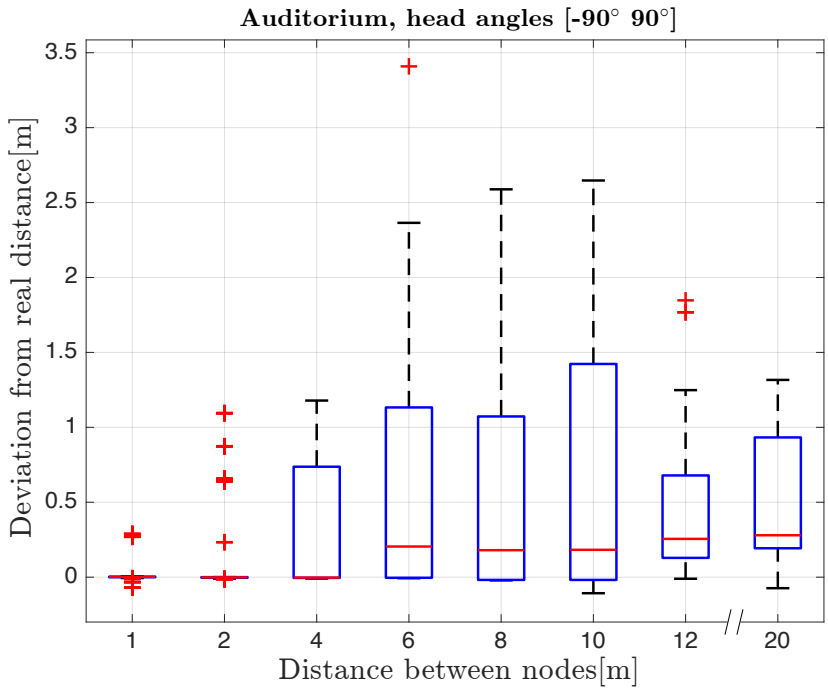


Figure 4.25: Distance calculation in auditorium, all measurements[a].

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

the distance deviation in millimeter. The standard deviation and variance can be seen in appendix B.2 for the two excerpts at respective distances. Thus, for proper detection, the system performance with relation to distance is excellent.

Figure 4.27 shows the impulse response captured at 20m distance in the auditorium. The measurement was performed on-axis with the nodes pointing directly towards each other. The figure shows that both channels were able to correlate the direct signal, with a strong reflection lagging approximately 1m behind on the right channel. The arbitrary reflections can also be seen.

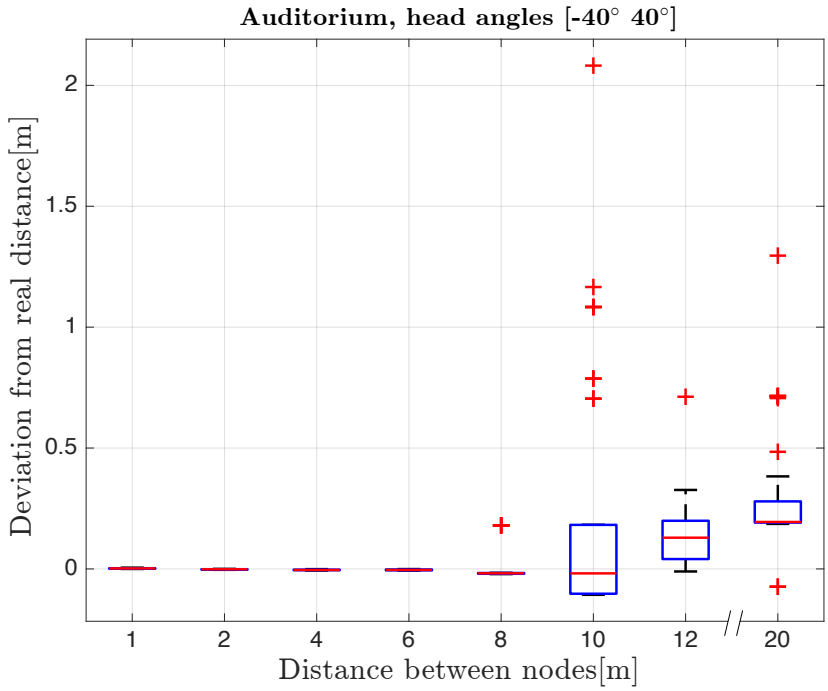


Figure 4.26: Calculated distances in auditorium for angles ±40°. For distances 1 to 8m, the boxes[a] representing the distribution are to small to be viewed, indicating small deviation. Figure 4.28 will represent the excerpt to more detail.

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

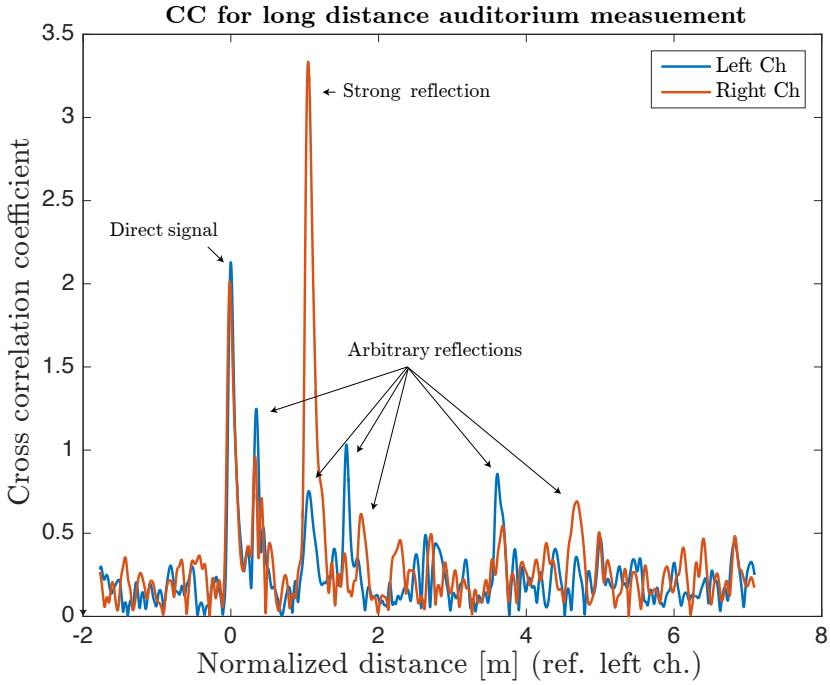


Figure 4.27: Impulse response for auditorium measurement at 20m distance.

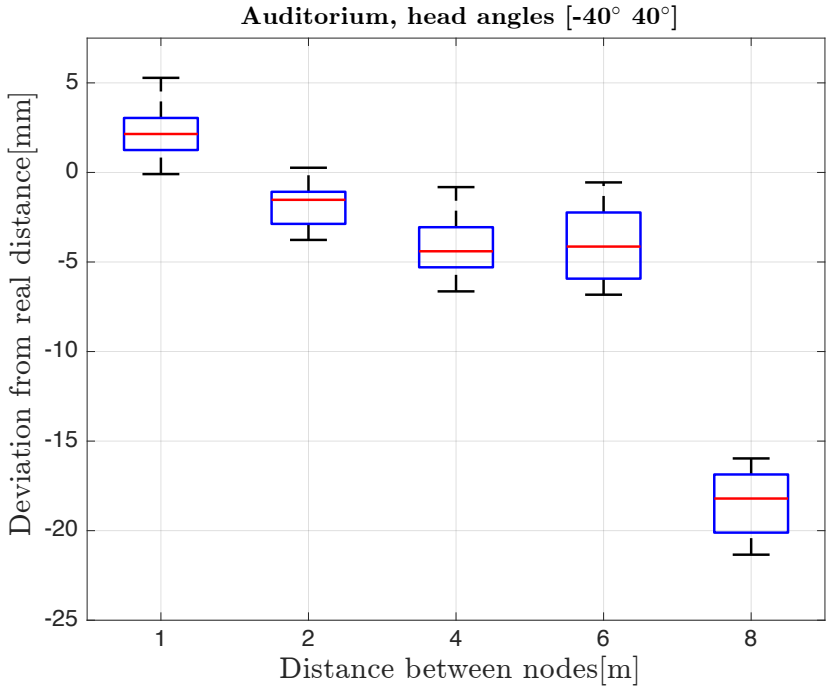


Figure 4.28: Calculated distances in auditorium for angles ±40°, 1 outlier omitted. Y-axis represented in [mm].

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

4.4 Discussion

This section provides a summary of the results from the measurements and evaluates the findings in relation to the system specifications presented in section 3.2. The main challenges from the results will be presented and discussed, as well as suggested improvements of the model and the implementation.

As mentioned in the introduction, the process of designing and implementing a functional model of a localisation system is a complex and challenging task. Many fields of expertise must be combined for the model to function in compliance with the desired specifications. A functional prototype system has been implemented in a limited amount of time, and many design choices have been made during the processes. The iterative approach revealed several improvements, but because of the limited time, many of them had to be left out. Therefore, suggestions regarding future improvements are provided at the end of this section.

4.4.1 System model and implementation

As mentioned earlier in the chapter, the method for detection of the pulse is rather primitive. The Carrier-synch demodulator approach includes a search strategy which took the first correlation peak above a certain threshold, in each buffer, as the arrival time of the pulse. This approach was not pursued further in the Complex demodulator implementation, as focus aimed at improving the accuracy and detection capabilities. In hindsight, this would have been a quick and efficient improvement, and significantly better results would be achieved in the measurements.

Other implementation improvements that has not been pursued is the decimation/symbol detection algorithm. In this work, it is based on hard filtering and decimation, but could have been improved with more sophisticated algorithms and an integrate-and-dump strategy. The threshold for signal detection could also have been improved. Adaptive regulation of a normalized coefficient based on the system logic may ensure that the system adapts to noise level.

The bug causing large deviation in the distance was not investigated thoroughly. However, indications point in the direction of absolute time mismatch, caused by a pure logic error.

The near-far problem is a limitation to the multi-user aspect of the system. The simulation showed excellent performance down to -12dB, but the prototype measured only capabilities of -5 to -5dB. The bug described at the beginning of this chapter is one plausible cause, another is the reflections present around the head. The transmitter frequency response may also contribute to the performance loss.

The complete DS-CDMA functionality has not been implemented. As discussed in Section 3.5.4, the preamble correlator should be capable of CDMA multi-user detection with only small logic changes. The implementation was not pursued, as time limited further implementation. Another important consideration is the fact that a preamble signal design not necessarily is the best choice for a real system.

The preamble strategy is a result of time accuracy problems. The unique sequences may possibly be correlated more efficiently in dedicated hardware such as FPGA's, thus, a full-rate correlation may be performed for each unique sequence without issues related to computational power. Another approach, possibly the best, is to investigate more sophisticated detection implementations. They do exist for RF-based communication systems and RADAR, but is based on complex signal processing and needs further research for use with burst transmission ultrasonic signals. These methods also mitigates Doppler issues, which is completely omitted in this work, and most likely causes the system to be unable of detection for moving nodes.

4.4.2 Angle accuracy

For evaluation of the accuracy, the standard deviation and variance are important measures. These variables describe the accuracy of the system when a series of measurements is performed in a similar situation. However, the accuracy must be regarded with respect to the purpose of the system. As discussed in section 3.2, the systems angle calculation is specified to operate under the constraints given by the perceptible differences of the human auditory system. Considering the results from the large quantity measurements shown in Figure 4.8, the deviation are very low, $\sigma_{deg} < 0.46$ and $\sigma^2_{deg} < 0.21$ for all angles. This indicates that the system implementation is capable of providing the required accuracy with respect to time resolution.

However, the angles acquired deviates from the real angle. As shown in Figure 4.8, the estimated angle does not correspond with the physical angle at a distance of 1m, with largest deviation of $-6°$ at $70°$. Deviation for all angles combined, $\sigma_{deg} = 2.4$ and $\sigma^2_{deg} = 5.7$, suggests that the actual precision will vary, causing abrupt changes in angle.

The reason for the deviation is attributed to reflections inflicted by the head. The figures comparing measurements with and without head, Figures 4.6, 4.7 and 4.11 verifies the influence of the head as a a reflecting surface. In general, the measurements performed with the head gave a stochastically fluctuating result around zero deviation, whereas measurements without the head results in a smoother curve.

It should be noted that the head used in the setup has a solid plastic surface and that no measurements was performed to quantify the reflective characteristics compared to a real head. As seen from the impulse response in Figure 4.13, the correlation from the reflection of the head was stronger than the direct signal at $80°$. It is plausible that this not will be the case if the system is worn by humans, due to absorption in skin and hair. This should be investigated for further work. As well, the implementation does not handle reflections well, as it assumes the largest correlation peak to be the direct pulse. Using a more advanced detection algorithm, the head reflection issue may be of less concern.

Another solution to the problem may be to position the microphones at the top of the head, thus, drastically reducing the influence of the head reflections as sug-

gested by the measurements. This will also avoid the problem of measuring the microphone spacing for each user, which will be necessary for a system with variable spacing due to varying head size.

The preliminary tests of the attenuation caused by the head shadow indicated a 50dB attenuation of sound pressure at $\approx 60°$ and virtually absent signal for angles larger than 75°. As such, it was expected that angle calculation would be challenging for the default setup, where microphones are placed on top of the ear muffs. The suspicion was confirmed by measurements, as angles above 70° was not detected. This aspect also underlines that a top-mounted configuration might be a future solution. Pros and cons of this alternative approach include

+ The microphones are visible for other nodes in 360° orientation, allowing for more accurate localization at angles above 70°, which was not possible for the default setup due to attenuation.
+ The calculation of angle will also be simplified. Since there is no obstructions the plane wave approximation in Equation 2.4 can be utilised for all incoming angles.
+ Increased range for angle calculation due to less attenuation of microphones.
+ The potential deviation caused by varying head size can be reduced if the microphones are mounted on a rigid base.
- The possibility to use the same microphones in a combined noise reduction system is left obsolete, thus requiring separate microphones for the respective features.
- Flexibility reduced if e.g. helmet is needed, specialized solution needs to be applied.

4.4.3 Distance accuracy

The results of the distance measurements performed in the anechoic chamber provided calculations with rather low deviation, $\sigma < 6.26 \cdot 10^{-4}$[m] and $\sigma^2 < 3.9 \cdot 10^{-7}$[m], for all angles[5]in the appendix. However, the calculated distance will vary with the angle of the head due to the fact that the calculated distance is the mean value of the two channels. This leads to a larger deviation for a given position. The expected distance accuracy presented in section 3.2 of ±1.8mm is not maintained for a given distance considering all angles. However for each respective angle in the measurement, the expected distance accuracy is maintained, see table B.4 in appendix for values. Regarding the purpose of the system, the distance accuracy is considered to be highly sufficient.

The distance measurements performed in the auditorium showed the systems vulnerability for reflections. Therefore, the most accurate results were attained for narrow angles(±40°) where both channels was able to correlate the direct signal path. For distances 1 to 6 meter, the standard deviation(σ) was less than 2mm.

[5] σ and $sigma^2$ for respective angles is found in Table B.4

For larger angles, the channel obstructed by the head tended to detect a reflected signal, rather than the direct signal. This resulted in a longer distance result. This behaviour was also quantified in the controlled reflection measurements. This issue may be solved by the implementation improvement mentioned earlier.

4.4.4 Error probability

As discussed in the measurements, a bug in the implementation randomly resulted in improbably long distance estimations. In the large quantity measurements, erroneous result occurred for seven measurements. The probability of error is thus $P(e) = 1.71 \times 10^{-3}$. These errors were of such a magnitude that they easily could be discarded as non-valid results in the implementation if more sophisticated logic is exploited. The angle calculations corresponding to the same measurements did not report deviating results. These random values also appeared in the measurements with continuous movement for 6 of the measurements, $P(e) = 6 \times 10^{-3}$.

As discussed in section 3.2, the link budget for the proposed system indicate a theoretical range of 7 to 30m depending on the background noise. Unfortunately, due to lack of suitable equipment, the background noise in the auditorium could not be verified, but it is assumed that the noise level was low. The longest distance tested in the measurement session was 20m, which resulted in a probability of detection error $P(e)_{\pm 90} = 0.22$, for all angles[6], whereas angles $\pm 40°$ resulted in no detection errors[7]. As discussed in section 3.2, the directivity of transmitters and receivers will affect the attainable range, and it was therefore expected that the steep angles would have some reduced probability for detection. It was not possible to verify the maximum predicted range due to space limitations. However, 20m detection range is considered as a good starting point for further work. It should also be noted that as discussed in section 3.2 the selected carrier frequency used in the implementation was not optimal for the applied transmitter. Figure 3.3 indicate that the relative SPL at the used carrier frequency is 15dB below the optimal frequency of the transmitter, longer range could therefore be expected by optimizing the use of the transmitter.

4.5 Feasibility

In general, the prototype performance regarding time precision is sufficient. The deviation in angular estimation related to the real angle is of greater concern, but may be improved by better detection algorithms and further investigation of microphone location. The distance measurements are far better than required for the system.

Considering these remarks, and the discussion above, the prototype indicates that further investigation of such a system may result in major improvements. Using

[6] 95 measurements in total
[7] 45 measurements

more sophisticated detection algorithms, the precision and reflection performance may be increased. Research indicates that Doppler-shift and other channel challenges may be mitigated. Shorter transmission times may also be achieved by using transducers with a wider band, and the noise sensitivity may be compensated with longer chip sequences. This will also help to reduce the near-far problem. A different PN-sequence may as well improve the SNR performance. If near-far shows to be a persisting problem, other advanced multiple access methods may be investigated, such as Frequency Hopping CDMA or TDMA using RF-synchronization.

This work proves the concept of a pure ultrasonic positioning system. However, as synchronization and logic may be moved to an RF-system, which indeed is necessary to transmit voice signals, the system can be designed such that less transmission is required in the ultrasonic channel. Thus, the refresh rate of the system may be increased, and a more advanced model based on the exchange of relative positions between nodes may be developed.

The inclusion of an IMU mitigates the problem of rapid change in angular position of the user. This data will be valuable in any of the future scenarios, and considering the availability of this technology, the authors may only see positive effects of the addition of such a system in future work.

Unfortunately, time did not permit a subjective evaluation of the system including HRTF processing, but the prototype is designed to easily allow for further investigations. The authors suggests that the HRTF processing is performed on separate computers, such that the workload of the prototype system is not exceeded.

4.6 Summary of improvements and future work

Below is a summary of the improvements suggested for future work.

Improvements of the physical model

- The reflective characteristics of the head used in the test setup should be investigated and compared to a real head, to quantify if the reflections leading to deviation in the calculated angle will apply to future implementation.
- To further investigate the problem with signals reflecting of off the head before reaching the microphone, measurements with lower angle increments may be performed.
- Measures to minimize the possibility of reflections from the head may be investigated. This includes accurate positioning of microphones in relation to the directional characteristics of the microphones, i.e. pointing the low sensitivity angle of the microphone towards the reflecting surface of the model.
- A known issue with the model is the front/back ambiguity. Signals arriving from behind will have the same relative time difference between microphones, leading to uncertainty if the signal is arriving from the front or back. Possible

solutions are using a third microphone in the front of the model, for instance a speech microphone, to determine direction.

- To properly compensate for the extended travel length, measurements with smaller angle increment should be performed to accurately map the increasing distance. Another approach may be to compensate by a correction factor as a function of angle, based on the head model.

Improvements of the prototype implementation

- Implementing a maximum allowable time difference between microphones would solve some of the errors discovered in the measurements. By using the first incident correlation as reference and only allowing detection at the opposite microphone for a limited time duration, many correlations caused by reflections can be discarded as illegal values. The maximum time duration should be limited by the time a sound wave uses around the head.
- The IMU in the test was merely for the proof-of-concept. Therefore, the IMU implementation must be fully investigated. Dedicated IMU's may use much higher refresh rates.
- The length of the chip sequence used in the system should be investigated further. Long codes obtain processing gain for robustness against noise and allows for more users. However, shorter codes provides faster correlation, and thus faster update time.
- The type of code used should be further evaluated. As discussed in section 2.4.3, other types of PN-sequences may provide better properties for the given application.
- Better detection methods should be considered. Doppler-shift may be corrected, and computational power requirements reduced.
- The preamble sequence should be considered. Better detection methods may eliminate the need for a preamble code.
- Other transducers should be considered to increase system performance. The selection addresses bandwidth, cost, sensitivity and output level.

Chapter 5

Conclusion

In this work, an ultrasonic positioning system for aiding a verbal communication in noisy environments has been proposed. A system model has been presented, the underlying technology has been described, and a prototype has been made. Finally, measurements and evaluation of the prototype has been presented.

The ultrasonic domain proves to be useful in short-distance positioning, specifically in the case of a speech enhancement system. Several advantages are evident, among them accuracy, privacy and independence from a centralised mainframe. However, the ultrasonic channel is limited in capacity, the background noise is highly dependent on environment, and includes many challenges like reflections, sound absorption and significant vulnerability to Doppler-shift. Local Positioning Systems has been investigated in the research literature, showing good performance for permanent installations.

Using technology originally developed for RF communication systems, the proposed system proves its function. Measurements show that the fairly simple prototype implementation performs well in relation to the positioning reproducibility and accuracy. However, challenges induced by the system model, specifically the head-worn hearing protector with transducers attached, implies that the accuracy of angular estimation is not sufficient for a fully functional system. By investigations of the cause of inaccuracy, this work suggests that the measures can be taken to mitigate the problem without compromising the intended design and function.

As the prototype system spans a wide range of speciality fields, a wide range of improvements are suggested. Many of those addresses communication technology, the basis of the positioning system. One of the most important in this regard is further development of the detector algorithm, to mitigate Doppler-shifts, increase SNR-capabilities, add proper multi-user access performance and increase the system refresh rate.

The location of transducers is a key design choice. The position suggested by the original system model will, as expected, introduce a range of challenges to the system. The results obtained by measurements indicates that a better performance

may be achieved with top-mounted microphones. However, if future studies indicate that head reflections is not as compromising as indicated by the measurements in this work, the originally suggested model may prove sufficient for the application.

As the speech transmission of the system depends on an RF-channel, the future system should consider an RF-based synchronization and logic communication channel. This may reduce the complexity of the ultrasonic positioning system itself, reduce channel access requirements, and as well provide further advantages such as location map exchange between several nodes. Thus, the accuracy and range of the system may prove to be increased. As well, near-far effects may be reduced or even eliminated with a TDMA channel access scheme. However, the results presented in this work indicates that a system based solely on ultrasound is feasible.

In general, the authors conclude that the proposed system based solely on ultrasonic technology is feasible for a full-scale implementation, provided that suggested improvements are pursued. The system will, in a future implementation, be able to precisely locate and position sources with a refresh rate sufficient to emulate a real spatial representation of verbal communication in noisy environments.

Bibliography

[1] S. Hodges J. Newman P. Steggles A. Ward M. Addlesee, R. Curwen and A. Hopper. Implementing a sentient computing system. *IEEE computer*, 2001.

[2] A. Chakraborty N. B. Priyantha and H. Balakrishnan. The cricket location-support system. *6th ACM International Conference on Mobile Computing and Networking*, 2000.

[3] Sonitor AS. URL `http://www.sonitor.com`. [Accessed: 10 Oct 2015].

[4] C. Randell and H. Muller. Low cost indoor positioning system.

[5] A. Hopper M. Hazas. Broadband ultrasonic location systems for improved indoor positioning. *IEEE Transactions on mobile computing*, 5(5), 2006.

[6] J. R. Gonzalez and C. J. Bleakley. High-precision robust broadband ultrasonic location and orientation estimation. *IEEE journal of selected topics in signal processing*, 3(5), 2009.

[7] J. Blauert. *Spatial Hearing- The Psychophysics of Human Sound Localization The Psychophysics of Human Sound Localization*. MIT Press, 1997.

[8] S. R. Price R. C. G. Smith. Modelling of human low frequency sound localization acuity demonstrates dominance of spatial variation of interaural time difference and suggests uniform just- noticeable differences in interaural time difference. *PLOS one*, 9, 2014.

[9] URL `https://en.wikipedia.org/wiki/Human_head`. [Accessed: 18 Dec 2015].

[10] B. D Simpson D.S Brungart, A. J Kordik. Effects of headtracker latency in virtual audio displays. *J. Audio Eng. Soc*, 54(1/2), 2006.

[11] A. Krokstad. *Akustikk for ingeniører*. NTNU, inst. for teleteknikk, 1999.

[12] A.B. Coppens J. V. Sanders L.E. Kinsler, A. R. Frey. *Fundamentals of acoustics*. Wiley & sons, 2000.

[13] A.J. Zuckerwar D. T. Blackstock H. E. Bass, L. C. Sutherland and D. M. Hester. Athmospheric absorption of sound: Further developments. *J. Acoust. Soc. Am*, 97(1), 1995.

[14] M. M. Saad T. Ballal and C. J. Bleakley. Design and implementation of an indoor ultrasonic communication system.

[15] S. Holm. Hybrid ultrasound–rfid indoor positioning: Combining the best of both worlds. IEEE, 2009.

[16] Jammet et.al. Interim guidelines on limits of human exposure to airborne ultrasound. *Health physics*, 46(4), 1984.

[17] E. D. Kaplan and C. Hegarty. *Understanding GPS: Principles and Applications-2nd Edition*. Artech House, Inc, 2006.

[18] J. Urena et. al. Advanced sensorial system for an acoustic lps. *Microprocessors and Microsystems*, 31, 2007.

[19] J.Guevaral J. L. Ealo Fernando A. Secol J. 0. Roal F. X. Ramos J. C. Prieto, A.R.Jimenez. Subcentimeter-accuracy localization through broadband acoustic transducers.

[20] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[21] Simon Haykin. *Communcation Systems*. John Wiley & Sons, Inc, 4th edition, 2001.

[22] A. Hernandez A. Jimenez D. Ruiz F. J. Alvarez C. D. Marziani M. C. Pérez, J. Urena. Performance comparison of different codes in an ultrasonic positioning system using ds-cdma. *IEEE International symposium on intelligent signal processing*, 2009.

[23] R. Gold. Optimal binary sequences for spread spectrum multiplexing. Technical report, IEEE Transactions on Information Theory, 1967.

[24] T. Kasami. Weight distribution formula for some class of cyclic codes. Technical report, University of Illinois, 1966.

[25] H. Boche S. Stanzak and M. Haardt. Are las codes a miracle. Technical report, Proc. IEEE Global Telecommunications Conference GLOBECOM'01, 2001.

[26] P. T. Gough. Signal processing and correlation techniques. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.2966&rep=rep1&type=pdf`, Assessed 30th November 2015. Department of Electrical and Computer Engineering, University of Canterbury.

[27] J. P. Costas. Synchronous communications. Technical report, Proceedings of the IRE 44 (12), 1556.

[28] M. Renfors. Synchronisation in digital receivers. [Accessed: 4 Nov 2015], 2014.

[29] K. W. Martin. Complex signal processing is not complex. *IEEE transactions on circuits and systems*, 51(9), 2004.

[30] F. J. Álvarez J. Lozano J. A. Paredes, T. Aguilera and J.Morera. Analysis of doppler effect on the pulse compression of different codes emitted by an ultrasonic lps. *Sensors*, 11(11), 2011.

[31] U. Mengali and M. Morelli. Data-aided frequency estimation for burst digital transmission. *IEEE Transactions on Communications*, 45:23–25, 1997. [Accessed: 15 Nov 2015].

[32] J. J. Garcia J. Urena E. Garcia, S. Holm. Link budget for low bandwidth and coded ultrasonic indoor location systems.

[33] S. Holm. Ultrasound positioning based on time-of-flight and signal strength. *International conference on indoor positioning and indoor navigation*, 2012.

[34] H. E. Bass. L. N. Bolen. Ultrasonic background noise in industrial environments. *J.acoust. Soc. Am.*, 78(6), 1985.

[35] S. Holm. Optimum fft-based frequency acquisition with application to cospas-sarsat. *IEEE transactions on aerospace and electronic systems*, 29(2), 1993.

[36] Real time audio processor matlab example script, 2012. URL `http://www.mathworks.com/matlabcentral/fileexchange/37770-real-time-audio-processor`. [Accessed: 4 Sep 2015].

[37] 2011. URL `http://www.mathworks.com/matlabcentral/fileexchange/32956-costas-loop/content/Costas_loop.m`. [Accessed: 12 Nov 2015].

Appendix A

Matlab class reference

The MATLAB code is structured as classes. The `RTAP` class is the fundamental class, containing the GUI, audio interface access and control of turntable. The GUI contains objects that call up other functions. Classes initialized by the `RTAP` is the `Node` class and the `Turntable` class. As well, the handle to a `mobiledev` class is stored in `RTAP` to provide access to the IMU data.

The `Node` class contains all the processing classes, and holds the relevant configuration proerties for the system.

Each of the processing classes are stand-alone, providing a constructor, a step function and diverse extra functions.

Each object is initialized upon creation of the class, such that all necessary parameters is set up only once. Thus, the parameters contained by the `Node` is passed on to each processing class.

After the creation, objects are interfaced by their `step()` function, which takes an appropriate input such as a signal or a vector or matrix. The output is a new signal or logic data.

A.1 Final implementation processing classes

`Modulator`

```
% Constructor:
Modulator(fs,fc,fm,code)

% <code> is a matrix of codes
% the preamble is generated separatly - omitting
% chip length inherent in <code>

% Step function:
signal = step()

% Function to assign output chip sequence
setOutput(code_idx, offset)
```

```
% <code_idx> is the index of the chip
% <offset> is the output start position in samples ref fs

% Function to modulate a code, simulation purpose
[s,m] = generateCode(code_idx)
% <s> is the modulated output
% <m> is the transposed sequence [-1,1]
% <code_idx> is the index of the chip

% Function to modulate an input signal, simulation purpose
[s,m] = modulate(pn)
% <s> is the modulated output
% <m> is the transposed sequence [-1,1]
% <pn> is the input sequence

% Continous ouput of preamble, SNR measurements
signal = continousOutput()
%<signal> is the modulated preamble
```

Detector2

```
% Constructor:
Detector2(fs, fc, fm, preamble, code)
% <preamble> is the transposed preamble chip
% <code> is the transposed chip matrix

% Step function:
detection = step(r)
% <detection> is the logic signal containing buffer lag
%       and detection threshold for each sequence
% <r> is the received signal
```

TimeKeeper

```
% Constructor:
TimeKeeper(n_code, fs, l_signal)
% <n_code> is the number of chips
% <l_signal> is the length of a modulated signal

% Step function:
[updates, current_lag] = step(detection)
% <detection> is the logic signal output from Detector2
% <updates> is a logic signal with updated codes
% <current_lag> is the absolute time in samples ref fs
```

Communicator

```
% Constructor:
Communicator(nodeId, fs, fc, fm, l_signal, code, n_code, ...
modulator)
% <nodeId> is the number
% <l_signal> is the length of a modulated signal
% <code> is the chip matrix
% <n_code> is the number of chips
% <modulator> is a handle to the modulator implementation

% Step function:
positions = step (updates, current_lag)
% <updates> is a logic signal from the TimeKeeper
% <current_lag> is a the absolute time from the TimeKeeper

% Send a request command
sendRequest()

% Calibrate at 1 metre
calibrate()
```

```
% Send periodic request commands
startPeriodicTransmission()

% Stop sending periodic request commands
stopPeriodicTransmission

% Property of interest
recorded_positions % Contains all detections observed
%       during periodic transmission
```

Positioner

```
% Constructor:
Positioner(n_node, handles)
% <n_node> is the number of nodes
% <handles> is a handle passed on by the RTAP
%       especially the mobildev handle for IMU

% Step function:
positions = step(USpositions)
% <positions> is the current positions of other nodes
% <USpositions> is a positions obtained by the
%       Communicator
% <current_lag> is the absolute time in samples ref fs
```

A.2 Additional functions

RTAP

```
% Constructor:
RTAP()

% Stores the handle of a mobildev object from command line
setMobdev(mobdev)
```

Node

```
% Constructor:
Node(nodeId, N_node, fs, fc, fm, handles)
% <nodeId> is ID defined in RTAP GUI
% <N_node> total number of nodes
% <handles> handles to properties in RTAP

% Step function:
out = step(in)
% <in> is the input signal from the RTAP buffer
% <out> is the ouput signal generated by Modulator
```

getGold

```
% returns the gold code matrix
[ code ] = getGold( start_index, n_codes, l_code)
% <code> is the matrix of gold codes
% <start_index> is the start code number
% <n_codes> is the number of codes
% <l_code> is the chip length
```

firDesign

```
% generates a LP filter object
[ firObject ] = firDesign( N,Fp,Fs,Rp,Rst )
% <firObject> is a matlab filter object
% <N> is the filter order
% <Fp> is the passband freq.
% <Fs> is the stop band freq.
% <Rp> is the pass band ripple
% <Rst> is the stop band ripple
```

A.3 Graphical user interface

The GUI is shown in figure A.1.

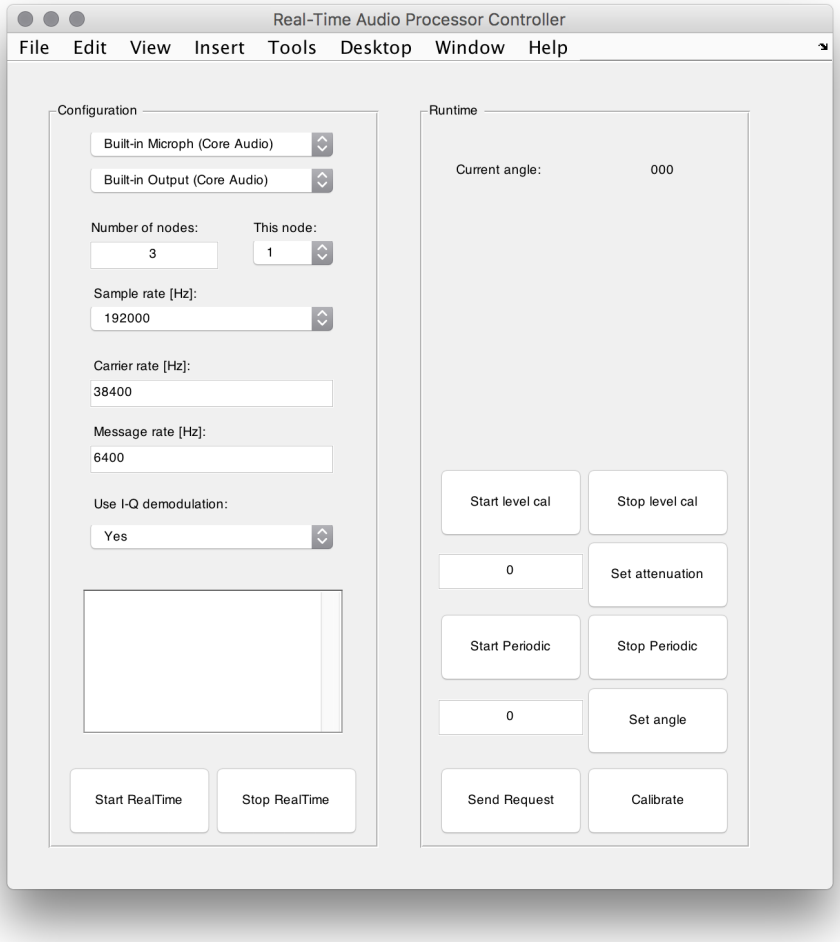


Figure A.1: MATLAB Graphical User Interface for node

Appendix B

Additional figures and tables

Table B.1: Equipment list with serial numbers

Type	Manufacturer	Model	Serial
Microphone capsule	Bruel&Kjær	4939	2546543
Microphone capsule	Bruel&Kjær	4939	2546544
Microphone capsule	Bruel&Kjær	4938	2496496
Microphone capsule	Bruel&Kjær	4938	2411398
Preamp	Norsonic	Nor1201	23824
Preamp	Norsonic	Nor1201	30490
Preamp	Norsonic	Nor1201	30517
Preamp	Norsonic	Nor1201	22038
Power supply	Norsonic	Nor336	20597
Power supply	Norsonic	Nor336	20626
Turntable	Norsonic	Nor265	29600
Audio Interface	Roland	Octa-capture	N/A
Audio Interface	Roland	Studio-capture	N/A
Amplifier	NAD	312	N/A
Amplifier	Acoustical Mfg. Co. Ltd	Quad 50E	CL4056
Audio interface	RME	Fireface 800	N/A
Laser distance meter	Bosch	PLR30	N/A
Hearing protector	3M Peltor	Optime III	N/A

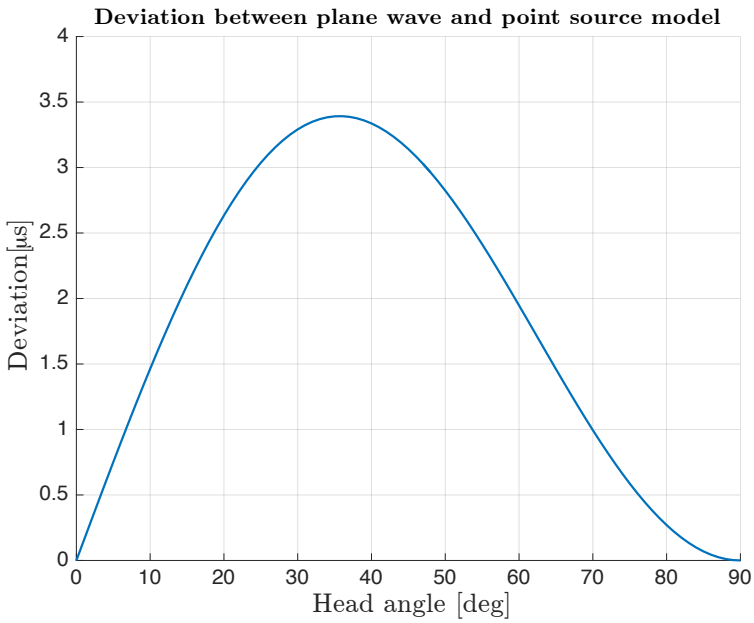


Figure B.1: Calculated difference in near and far field model from distance 3x microphone diameter.

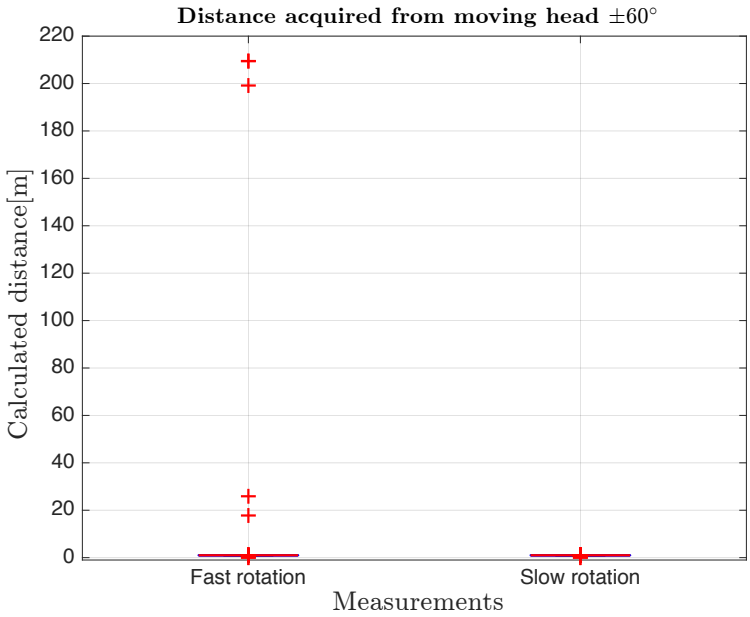


Figure B.2: Distance calculation for head panning(±60°) in the horizontal plane, omitting outliers[a].

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + 1.5 · boxrange, or lower than 25% percentile- 1.5 · boxrange.

Table B.2: Values for distance calculation in auditorium

	±40°			±90°			
Dist[m]	μ[m]	σ[m]	σ^2 [m]	μ[m]	σ[m]	σ^2 [m]	P(error)
1	1.002	0.0013	$1.8 \cdot 10^{-6}$	1.025	0.089	0.008	0
2	1.998	0.0011	$1.3 \cdot 10^{-6}$	2.175	0.347	0.121	0.011
4	3.996	0.0015	$2.39 \cdot 10^{-6}$	4.330	0.426	0.181	0.011
6	5.996	0.0021	$4.56 \cdot 10^{-6}$	6.513	0.716	0.513	0.011
8	8.003	0.063	$4.0 \cdot 10^{-3}$	8.609	0.769	0.592	0
10	10.195	0.478	0.228	10.697	0.791	0.626	0.042
12	12.132	0.132	0.017	12.417	0.449	0.2023	0.168
20	20.286	0.24	0.057	20.519	0.40	0.159	0.221

Table B.3: Large quantity measurements in anechoic chamber. Standard deviation and variance in sample difference between channels for angles [0°70°]

Angle	0°	10°	20°	30°	40°	50°	60°	70°
σ_{sample}	0.4334	0.4733	0.496	0.45	0.36	0.38	0.36	0.61
σ^2_{sample}	0.19	0.22	0.25	0.199	0.13	0.14	0.13	0.38

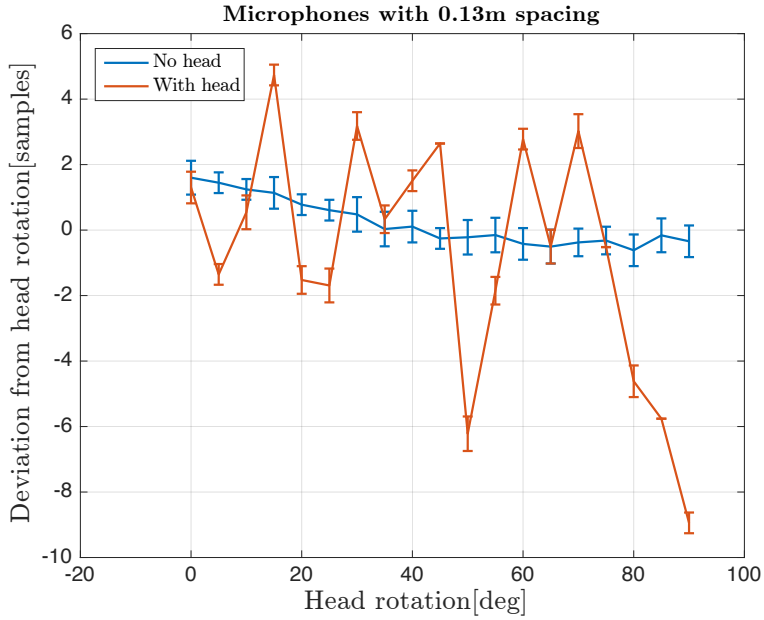


Figure B.3: Difference between calculated angle and real angle, presented in sample differnce, maximum $\sigma_{sample} = 0.527$

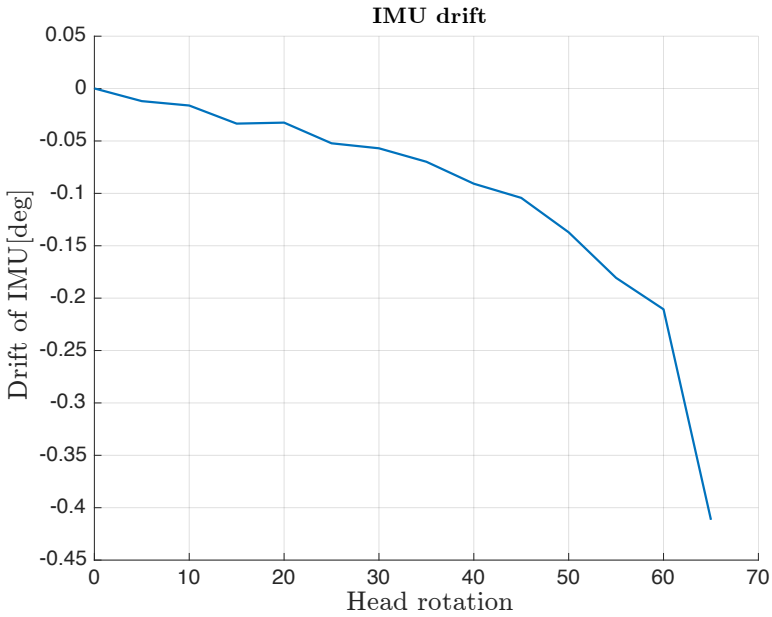


Figure B.4: IMU drift as a function of the head rotation.

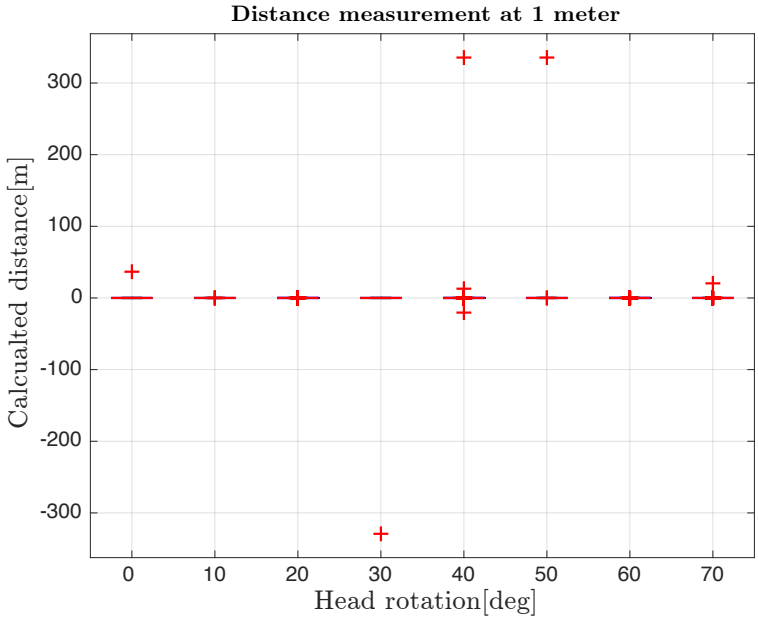


Figure B.5: Distance calculation performed at 1m, 500 measurements for each angle[a]. Indicating the position of the arbitrary outliers caused by a bug in the implementation.

[a]The red line represents the median, the blue box represents values inside the 25% and 75% percentiles. Outliers are plotted as individual points. Whiskers include the datapoints that is not considered to be outliers. Points are considered outliers for values larger than the 75% percentile value + $1.5 \cdot$ boxrange, or lower than 25% percentile- $1.5 \cdot$ boxrange.

Table B.4: Standard deviation and variance of the distance calculation for angles 0° to 70°. μ_{dev} designates the deviation from the real distance.

Angle	0°	10°	20°	30°
μ_{dev}[mm]	0.66	1.83	0.98	5.05
σ [mm]	0.38	$3 \cdot 0.1$	0.48	0.39
σ^2 [mm]	$1.4 \cdot 10^{-4}$	$9.0 \cdot 10^{-5}$	$2.34 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$
Angle	40°	50°	60°	70°
μ_{dev}[mm]	8.61	10.0	12.8	14.07
σ [mm]	0.24	0.3.5	0.3	0.6
σ^2 [mm]	$5.9 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$	$8.9 \cdot 10^{-5}$	$3.9 \cdot 10^{-4}$

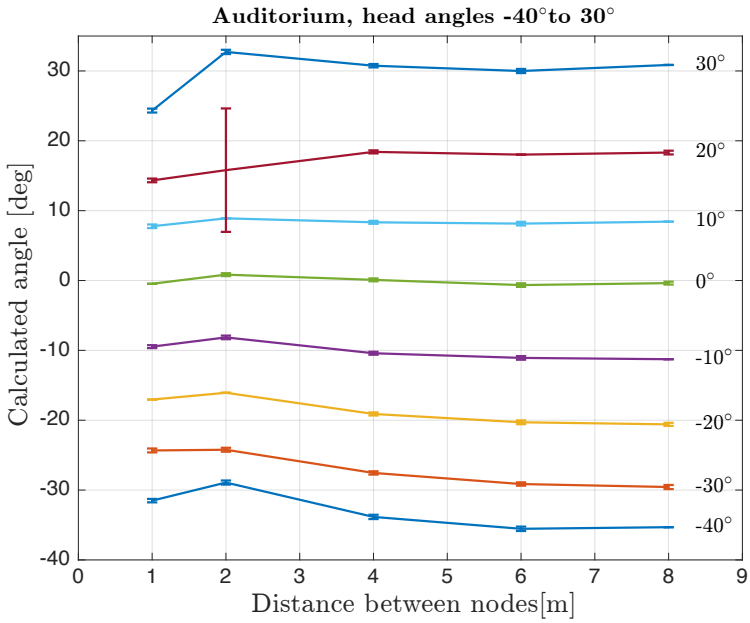


Figure B.6: Calculated angle for auditorium measurements as a function of distance. Datapoints show the mean of measurements with the standard deviation(σ_{deg}) as bars in each point

Figure B.7: Default test setup, microphones placed above each ear muff. Ultrasonic transmitter placed on top of head.

Figure B.8: Measurement setup in anechoic chamber.

Figure B.9: Measurement setup in auditorium.

Figure B.10: Measurement setup in auditorium.

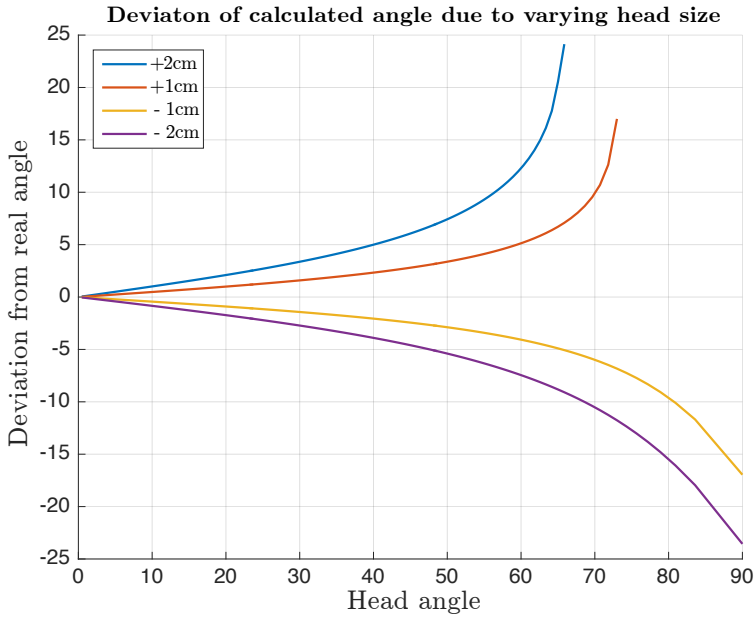


Figure B.11: Deviation of the calculated angle due to head size. $\pm$ indicates difference in diameter(d) from the average head size used in the implementation. For wider heads the travel will be longer, thus resulting in complex values before maximum angle reached(+1 and +2 graphs).

B.1 Calculation of angle adaption

As discussed in section 2.3.1, no simple equation exists to compensate for the extended travel length caused by the obstructing head. Therefore to be able to utilise the plane wave equation[2.4] with the extended travel caused by the obstructing head, a solution was suggested to compensate for the extended travel by adding a linear term to the plane wave equation.

The procedure was performed in the following procedure

- Perform measurements, acquiring the time difference between microphones, as a function of the real head rotation angle.
- Estimate the maximum time difference between microphones of the test model, which characterises the time difference at 90°.
- Find the point where the line-of-sight between transmitter and receiver ends, and thereby from where the linear term should be added to the calculation.
- calculate linear term
- Fit compensated calculation with test measurement by adjusting the parameters.

The linear term was simply derived from the common form of linear equation with two variables

$$y = mx + b \tag{B.1}$$

Whereas the constant m designates the gradient of the line, and b designates the crossing of the y-axis.

The linear equation written on the two-point form in the (x,y) plane can be written as

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) \tag{B.2}$$

rearranging and solving for y gives

$$y = x\frac{y_1 - y_0}{x_1 - x_0} + \frac{x_1 y_0 - x_0 y_1}{x_1 - x_0} \tag{B.3}$$

this gives us the gradient m;

$$m = \frac{y_1 - y_0}{x_1 - x_0} \tag{B.4}$$

and zero crossing point b

$$b = \frac{x_1 y_0 - x_0 y_1}{x_1 - x_0} \tag{B.5}$$

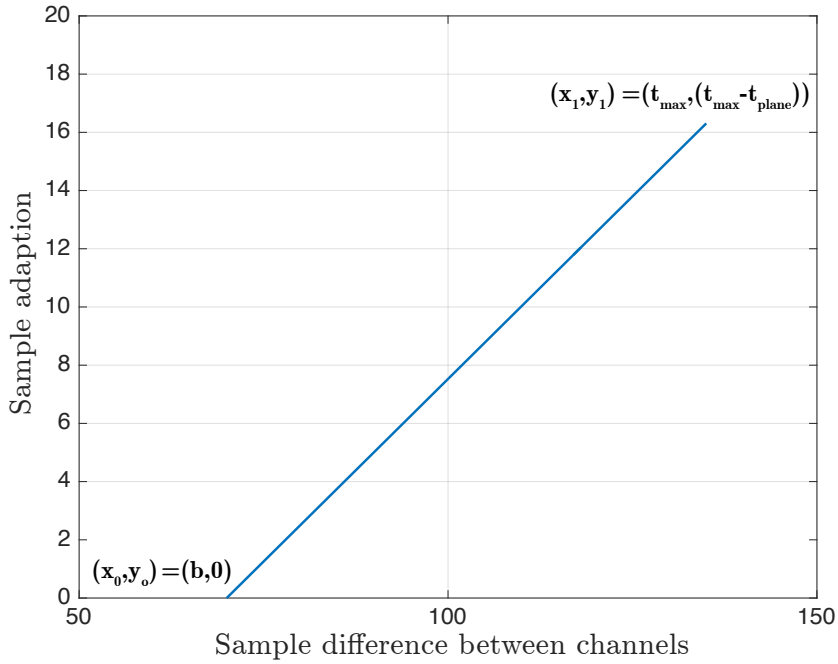


Figure B.12: Resulting linear term with example values in samples

Given the four known points in the plane, as seen in figure B.12

1. the breakpoint(b) from where the term should be applied, $y_0 = 0$
2. the maximum time difference from the plane wave equation[2.4], t_{plane}, which can be calculated from the plane wave equation by inserting the known distance between microphones(d).
3. the maximum possible difference(t_{max}) between channels, obtained from measurements at 90° angle.

Inserting the known values gives

$$f(\Delta t) = \Delta t \cdot \frac{t_{max} - t_{plane}}{t_{max} - b} - \frac{b \cdot (t_{max} - t_{plane})}{t_{max} - b} \tag{B.6}$$

Where Δt characterizes the measured time difference between channels. For the implementation the calculation needs to be divided in two parts. One for angles where line of sight between channels are maintained, whereas no adapting is performed. For angles larger than line-of-sight the linear term is added to compensate for the extended travel time due to obstructing head. The following equation represents the calculation of angles larger than line-of-sight:

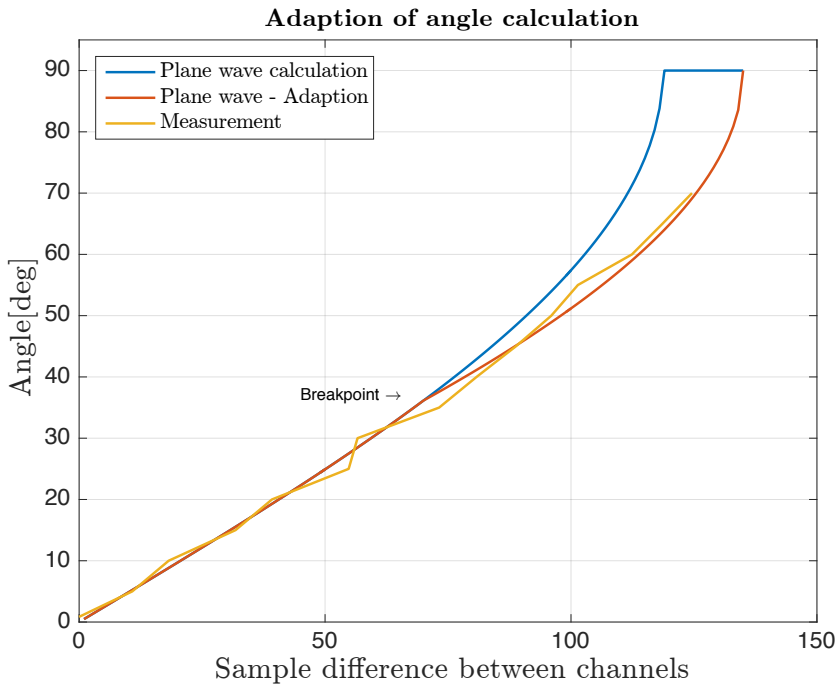


Figure B.13: Plane wave calculated

$$\theta = sin^{-1}\left(\frac{c_s}{d}\cdot\left(\Delta t - \Delta t\cdot\frac{t_{max}-t_{plane}}{t_{max}-b} - \frac{b\cdot(t_{max}-t_{plane})}{t_{max}-b}\right)\right) \tag{B.7}$$

The resulting angle calculation as can be seen in figure B.13, indicates that the adapted calculation corresponds slightly better with the measurements. It should however be appointed that this procedure was performed strictly as a proof of concept. For further testing more measurements should be performed with smaller incrementation and measures to avoid reflections from head should be taken.